Kwantowa kosmologia – to tam, gdzie fizyka rodzi się z metafizyki.

MICHAŁ HELLER

Claus Kiefer

KWANTOWY KOSMOS

Od wczesnego świata do rozszerzającego się uniwersum

Copernicus Center PRESS

Claus Kiefer

KWANTOWY KOSMOS

Od wczesnego świata
do rozszerzającego się uniwersum

tłumaczenie
Michał Tęcza

Copernicus Center PRESS

Spis treści

Karta redakcyjna

Dedykacja

Motto

Przedmowa do polskiego wydania

Przedmowa

1. Przestrzeń, czas, teoria względności

2. Teoria kwantowa

3. Entropia

4. Kosmologia

5. Grawitacja kwantowa

6. Funkcja falowa wszechświata

7. Refleksje

Uwagi

Literatura uzupełniająca

© Copyright by Copernicus Center Press, 2018
Copyright © S. Fischer Verlag GmbH, Frankfurt am Main 2008

Tytuł oryginalny
Der Quantenkosmos. Von der zeitlosen Welt zum expandierenden Universum

Adiustacja i korekta
Mirosław Ruszkiewicz

Projekt okładki
Michał Duława

Grafika na okładce
Alexey Brin, Naeblys | Fotolia.com

Skład
MELES-DESIGN

ISBN 978-83-7886-236-9

Wydanie I
Kraków 2018

Copernicus Center Press Sp. z o.o.
pl. Szczepański 8, 31-011 Kraków
tel./fax (+48 12) 430 63 00
e-mail: marketing@ccpress.pl
Księgarnia internetowa: http://ccpress.pl

Konwersja: eLitera s.c.

Dla Claudii

Każda nauka rośnie jak drzewo,
nie tylko gałęzie się rozrastają,
ale i korzenie wnikają głębiej.

David Hilbert

Przedmowa do polskiego wydania

To dla mnie wielki zaszczyt, że moja książka została przetłumaczona na język polski. Mam nadzieję, że również w Polsce znajdzie ona wielu zainteresowanych czytelników. Serdeczne podziękowania składam panu profesorowi Michałowi Hellerowi za jego zaangażowanie, jak również tłumaczowi oraz współpracownikom Copernicus Center Press za ich trud włożony w przygotowanie tego wydania książki.

Upłynęło już niemal dziesięć lat od pierwszej publikacji *Kwantowego kosmosu*, mimo to zawarte w nim tematy także dziś przedstawiłbym w ten sam sposób. Fascynujące poszukiwania grawitacji kwantowej trwają nadal. W ostatnich latach dokonał się jednak pewien postęp w obserwacjach i eksperymentach, co w niniejszym tekście nie mogło zostać uwzględnione. Kilka przykładów przytoczę poniżej.

Teleskop kosmiczny Plancka rozpoczął pracę w 2009 roku. Dzięki niemu przeprowadzono jak dotąd najdokładniejsze pomiary kosmicznego promieniowania tła. Zmierzone parametry kosmiczne wykazują tylko niewielką różnicę w porównaniu z pomiarami sondy WMAP. Satelitę wyłączono w 2013 roku.

Wspomniana w tekście cząstka Higgsa, tak istotna dla zrozumienia modelu standardowego fizyki cząstek elementarnych, została tymczasem odkryta w akceleratorze LHC w Genewie. Dowód na to opublikowano w 2012 roku. Dzięki temu z dużą precyzją potwierdzono model standardowy, jednak nie znaleziono dowodów potwierdzających inne teorie fizyczne (na przykład teorię supersymetrii).

Ogromnym postępem eksperymentalnym jest też pierwszy bezpośredni dowód na istnienie fal grawitacyjnych. Uzyskano go dzięki dwóm interferometrom LIGO, zlokalizowanym na terenie Stanów Zjednoczonych. W pracach nad technologią wykorzystaną w tych

urządzeniach istotną rolę odegrała grupa skupiona wokół interferometru GEO600 w okolicy Hanoweru. Wynik zarejestrowano 14 września 2015 roku i opublikowano w lutym 2016 roku. Odkryty sygnał powstał wskutek połączenia się dwóch czarnych dziur w odległości około 1,3 miliarda lat świetlnych, a energia wyemitowana podczas tego połączenia odpowiadała około trzem masom Słońca. To jak dotąd najpoważniejszy dowód na istnienie czarnych dziur. Od tej pory zaobserwowano kilka innych zdarzeń, do których doszło w wyniku połączenia się czarnych dziur. Niedawno zarejestrowano również zdarzenie, którego źródłem było połączenie się dwóch gwiazd neutronowych. Podczas niego udało się nawet zaobserwować rozbłysk gamma w widmie elektromagnetycznym. W 2017 roku Nagrodę Nobla za uzyskanie bezpośredniego dowodu na istnienie fal grawitacyjnych przyznano trzem najbardziej zaangażowanym w ten projekt naukowcom. Jesteśmy zatem świadkami narodzin nowej gałęzi astronomii. Wspomniany w tekście planowany projekt satelitów LISA, mający na celu zdobycie dowodu na istnienie fal grawitacyjnych, opóźnił się i nie wystartuje prawdopodobnie przed 2034 rokiem. Sonda przygotowawcza w ramach projektu LISA, tak zwana LISA Pathfinder, została już jednak z powodzeniem wysłana w kosmos w 2015 roku.

Postępy eksperymentalne widać także w obszarze mechaniki kwantowej, a zwłaszcza informacji kwantowej. Za pomocą satelity uzyskano stan splątany na odległość 1200 kilometrów. Splątane fotony zostały wykorzystane w celu utworzenia zabezpieczonego przed podsłuchem łącza do transferu danych między Wiedniem a Pekinem. W przygotowaniu są też eksperymenty, których celem będzie potwierdzenie bezpośredniej superpozycji skwantyzowanego pola grawitacyjnego. Kwantowy kosmos pozostaje nadal fascynujący!

Kolonia, listopad 2017
Claus Kiefer

Przedmowa

Czy można pojąć wszechświat? Czy można wyjaśnić jego powstanie i rozwój? Czy możemy formułować wiarygodne teorie na temat jego przyszłości? To pytania tak stare, jak sama kultura ludzka. Niniejsza książka postara się o odpowiedzi, których dostarczyć mogą nam nowoczesne nauki przyrodnicze. Ich siła opiera się na dwóch filarach: opisie teoretycznym i eksperymencie. Dopiero ta mieszanka sprawia, że jesteśmy w stanie wypracować spójny obraz natury. Stwierdzamy, że świat opiera się na prawach natury i dzięki temu staje się pojmowalny.

Współczesna fizyka została ukształtowana w dużej mierze przez dwie teorie: teorię względności oraz teorię kwantową. Teoria względności zmieniła sposób postrzegania przez nas przestrzeni i czasu, ponieważ grawitacja jest niczym innym jak geometrią czterowymiarowej czasoprzestrzeni. Teoria kwantowa nauczyła nas, że klasyczny obraz cząstek i ich orbit w obrazie mikroskopowym przestał już obowiązywać. Obie teorie zostały z powodzeniem potwierdzone dzięki eksperymentom, jednak nie można ich zastosować równocześnie dokładnie w takim samym stopniu. Jest to widoczne na przykładzie pojęcia czasu, gdyż odgrywa on w nich zupełnie inną rolę. Jedno z najważniejszych zadań w fizyce polega na opracowaniu teorii, która połączyłaby teorię względności i teorię kwantową. Taka teoria nazywana jest grawitacją kwantową.

Grawitacja kwantowa powinna umożliwić opisanie rzeczy, których w tej chwili nie można jeszcze wyjaśnić. Należą do nich na przykład powstanie świata czy los czarnych dziur. Ma ona nas też doprowadzić do nowych odkryć w zakresie mikroskopowej natury przestrzeni i czasu. Być może dzięki niej powstanie nawet teoria wszystkiego, na podstawie której przynajmniej w teorii uda się wyprowadzić wszystkie inne prawa natury.

Moja książka mówi przede wszystkim o poszukiwaniu grawitacji kwantowej i o jej zastosowaniu w odniesieniu do wszechświata jako całości, czyli o kosmosie kwantowym. Naturalnie podróż ta musi rozpocząć się wyczerpującym opisem teorii względności i teorii kwantowej oraz ich najważniejszych osiągnięć. Ponadto nie wolno zapomnieć o dyskusji na temat interpretacji teorii względności, która rozgorzała przy okazji słynnego i kontrowersyjnego przykładu z udziałem kota Schrödingera. Część 3 poświęcona jest centralnemu pojęciu w fizyce, czyli entropii. Wzrost entropii wyraża tak istotną dla nas różnicę między przeszłością a przyszłością. Teoria względności i teoria kwantowa umożliwiają również opisanie wszechświata, co prowadzi nas w części 4 do tematu kosmologii, jednego z najbardziej fascynujących obszarów fizyki, w ramach którego teorie nadal potwierdzane są dzięki licznym eksperymentom i obserwacjom. Część 5 to właściwe roztrząsanie tematu grawitacji kwantowej. Na początku uzasadnię, dlaczego taka teoria jest niezbędna, a później przedstawię jej najważniejsze eksperymenty: kwantową teorię względności, do której należą między innymi geometrodynamika i pętlowa grawitacja, a także teorię strun. Części 6 i 7 zostały poświęcone zastosowaniu grawitacji kwantowej w odniesieniu do samego wszechświata i fascynującym wnioskom, które z niej wynikają dla naszego obrazu świata i miejsca człowieka w świecie.

Oczywiście prezentowane tematy nie są łatwe w odbiorze, dlatego też każdy autor konfrontuje się z problemem poziomu szczegółowości ich rozpatrywania. Próbowałem wziąć sobie do serca radę Alberta Einsteina, zgodnie z którą dyskusję należy prowadzić w sposób tak prosty, jak tylko się da, ale nie prościej. Zrezygnowałem zatem w znacznym stopniu z wzorów, a przy przedstawieniu tematów starałem się na tyle uprościć treść, na ile było to możliwe bez jej zafałszowania. Przy czym po cichu założyłem, że niektóre podstawowe pojęcia z zakresu fizyki, jak na przykład masa, energia, pęd czy siła, są czytelnikowi znane ze szkolnych lekcji.

Fizyki nie da się dziś uprawiać w zupełnej izolacji, dlatego chcę wyrazić swoją wdzięczność wszystkim kolegom i studentom, z którymi w ciągu ostatnich 25 lat dyskutowałem na te tematy i z którymi prowadziłem badania. Tych osób jest zbyt wiele, by móc wymienić wszystkie z nazwiska, jednakże pragnę w szczególności podziękować prof. H.-Dieterowi Zehowi, który wprowadził mnie w ten świat, a z którego sposobu rozumienia fizyki wiele wyniosłem. Imienne podziękowania należą się także mojemu koledze, prof. Klausowi Volkertowi, który podjął się zadania uważnego przeczytania manuskryptu mojej książki i opatrzenia go wieloma pomocnymi komentarzami. Na końcu dziękuję również panu Gerhardowi Kollandowi, fizykowi dyplomowanemu, który z wielką pieczołowitością wykonał niektóre rysunki.

Kolonia, lipiec
Claus Kiefer

1. Przestrzeń, czas, teoria względności

By pojąć, że niebo wszędzie jest niebieskie,
nie trzeba podróżować po świecie.

Johann Wolfgang von Goethe

Od starego kosmosu do przestrzeni i czasu

W *Wyznaniach hochsztaplera Feliksa Krulla* Tomasza Manna profesor Katschka, podczas podróży pociągiem z Paryża do Lizbony, opowiada rzekomemu markizowi de Venosta o historii i przyszłości wszechświata: „Nie istniał bowiem byt zawsze i nie zawsze też będzie istniał. Miał on swój początek i będzie miał swój kres, a razem z nim również przestrzeń i czas, co tylko poprzez byt istnieją i tylko w bycie wiążą się wzajemnie" (wg przekładu Andrzeja Rybickiego, 1968). W tej inspirującej pogawędce obu panów w wagonie restauracyjnym chodzi właśnie o pytanie o wszechświat jako całość i jego rozwój w czasie. Od tysiącleci ludzkość zajmuje się tą kwestią, a jej wyjaśnienia bywały już bardzo odmienne. Niniejsza książka opisuje nasze nowoczesne, ukształtowane przez nauki przyrodnicze pojmowanie wszechświata, a także to, co wydaje nam się, że wiemy, oraz kierunek, w jakim nasza wiedza prawdopodobnie będzie się rozwijać. Czas, jaki przeznaczymy na naszą podróż, będzie porównywalny z długością dawnej podróży pociągiem z Paryża do Lizbony. Wyjaśnienia profesora Katschki są bardzo bliskie dzisiejszym przekonaniom, lecz nie obejmują jednego ważnego aspektu, a mianowicie roli teorii kwantowej. Stanie się to jasne dopiero pod koniec naszej podróży.

Naukę o powstaniu i rozwoju wszechświata nazywamy kosmologią. Greckie słowo „kosmos" oznaczało pierwotnie porządek bądź ozdoby – to

ostatnie znaczenie widoczne jest jeszcze w słowie „kosmetyka”, czyli sztuka zdobienia. Stare znaczenie celnie charakteryzuje antyczny obraz uporządkowanego świata, pełnego piękna i harmonii. Taki obraz świata przetrwał do XVII wieku, gdy wraz z początkiem nowoczesnych nauk przyrodniczych zaczęło się nieuchronne burzenie starego kosmosu. Historyk nauki Alexandre Koyré mówi o zburzeniu starego świata jako o głębokim, fundamentalnym procesie prowadzącym do tego, że człowiek utracił świat, w którym żył i nad którym się zastanawiał, i nie tylko musiał zmienić oraz ukształtować na nowo fundamentalne pojęcia i właściwości, lecz także sposób myślenia. Proces przemiany, pod kątem historycznym jeszcze stosunkowo młody, doprowadził do koncepcji rozszerzającego się wszechświata o ogromnych rozmiarach, który kiedyś będzie miał swój koniec i na tle którego Ziemia wraz z człowiekiem jest tylko nic nieznaczącym ziarnkiem piasku. Proces ten prawdopodobnie wciąż trwa, a my przekonamy się później, dlaczego tak się dzieje.

Na początek spójrzmy jednak na stary, utrzymujący się przez ponad dwa tysiące lat obraz świata. Wpłynął na niego kosmos Arystotelesa (384–322 r. p.n.e.), w którym panuje wyraźne rozdzielenie niebiańskiego i ziemskiego świata. W ziemskim świecie każda zmiana wynika z mieszania się lub rozdzielania czterech żywiołów, które wprowadził już Empedokles (około 485–425 r. p.n.e.): ognia, ziemi, powietrza i wody. Według Platona (427–347 r. p.n.e.) czterem żywiołom przyporządkowane są cztery z pięciu wielościanów foremnych. Są to ciała, których powierzchnie ograniczające stanowią regularne, przystające wielokąty: składający się z czterech trójkątów równobocznych czworościan odpowiada ogniowi, kostka składająca się z sześciu kwadratów – ziemi, zbudowany z ośmiu równobocznych trójkątów ośmiościan – powietrzu, i dwudziestościan powstały z dwudziestu równobocznych trójkątów odpowiada wodzie. Arystoteles odróżnia ruchy naturalne od wymuszonych. Ruchem naturalnym jest ruch prostopadły, wychodzący od Ziemi lub skierowany ku niej, w zależności od tego, czy ciało jest ciężkie (jak woda), czy lekkie (jak ogień). Tu po raz pierwszy pojawia się pojęcie grawitacji (ciężkości), choć jedynie w obszarze ziemskim.

Zupełnie inaczej dzieje się w świecie niebiańskim: tam istnieją jedynie ruchy naturalne i są to ruchy obrotowe bądź ruchy składające się z takich ruchów. Nieruchoma, kulista Ziemia jest jak cebula, otoczona ponad pięćdziesięcioma kulistymi powłokami (sferami), poruszającymi się po okręgu, z czego jedynie siedem z nich dźwiga ciała niebieskie: Księżyc, Słońce i pięć znanych starożytnym planet – Merkurego, Wenus, Marsa, Jowisza i Saturna. Pozostałe sfery są konieczne do wyjaśnienia nieregularnych ruchów tych ciał niebieskich. Do zewnętrznej sfery przyczepione są gwiazdy stałe. Materię, z której składa się świat niebiański, nazywano eterem lub *kwintesencją*, piątą esencją. W platonicznym przyporządkowaniu odpowiada mu składający się z dwunastu regularnych pięciokątów dodekaedr, ostatni z pięciu regularnych wielościanów. Nie istniały inne możliwości, ponieważ już w antyku można było matematycznie udowodnić, że istnieje jedynie pięć ciał tego rodzaju. W części 4 zobaczymy, jak w nowoczesnej kosmologii kwintesencja święci swój triumfalny powrót jako ciemna energia.

Arystotelesowski świat jest wieczny i nieprzemijalny. Jest także skończony – poza sferą gwiazd stałych nie istnieje dosłownie nic, nawet pusta przestrzeń. Z tego też powodu nie do pomyślenia było wyobrażenie próżni, gdyż mogłaby ona sugerować potencjalną nieskończoność świata. Według Arystotelesa spadające ciało mogłoby osiągnąć nieskończenie wielką prędkość, gdyby udało się doprowadzić do całkowitego zniknięcia hamującego oporu powietrza. Z uwagi na to, że było to niewyobrażalne, mówiono o *horror vacui* natury, czyli lęku przed próżnią. Nieskończoności znane były właściwie z geometrii, którą podsumował działający w Aleksandrii Euklides (od około 365 do około 300 r. p.n.e.) w swoim sławnym dziele *Elementy* i którą do dziś nazywa się na jego cześć geometrią euklidesową. Proste są tam nieskończenie długimi liniami, a proste równoległe nigdy się nie przecinają. Jednak świat matematyki był zupełnie innym światem, który nie miał nic wspólnego z kosmosem Arystotelesa.

Arystotelesowski obraz świata dominował aż do XVII wieku. Naturalnie istniały także inne mieszane wyobrażenia, gdyż musiano brać pod uwagę forsowaną przez religię chrześcijańską koncepcję stworzenia świata, w praktyce jednak nie były one zbyt istotne. Dlaczego ten obraz świata utrzymywał się tak długo? Główną przyczyną była być może głęboko zakorzeniona wiara w skończoność świata i poznania. Według Hansa Blumenberga to, że na świecie mogło istnieć coś, co dla człowieka jest nie tylko chwilowo czy przejściowo, lecz definitywnie niedostępne i niewidoczne dla jego naturalnych predyspozycji poznawczych, stanowiło w czasach antyku, jak i średniowiecza nieznaną, a przy określonych założeniach metafizycznych również niewykonalną insynuację. W pewnym sensie wszystko na świecie było już „zbadane" i wystarczyło powołać się na jakiegoś mędrca, takiego jak Arystoteles, który wszystko to pieczołowicie sprawdził i spisał. Kolejną rolę odgrywał rozdział świata niebiańskiego i ziemskiego, zwanego światem sublunarnym. Nie istniała fizyczna rzeczywistość poza zjawiskami niebieskimi, nikt nie był zainteresowany właściwościami gwiazd. Astronomia była bardziej sztuką niż nauką. Chodziło tylko o to, aby opisać i przewidzieć ruchy ciał niebieskich, dlatego też tytuł głównego dzieła słynnego aleksandryjskiego astronoma Klaudiusza Ptolemeusza (od około 90 do około 160 r. n.e.) znaczył tyle co *Matematyczny zbiór* (dzieło znane bardziej pod arabskim tytułem *Almagest*), ponieważ Ptolemeusz był zainteresowany jedynie formalnym opisem i nie musiał zajmować się eteryczną substancją sfer niebieskich.

Zwyczajowo zburzenie starego kosmosu przypisuje się Mikołajowi Kopernikowi (1473–1543), czemu daje wyraz często stosowana metafora o przewrocie kopernikańskim. Obraz świata według Kopernika był w istocie tylko pierwszym, choć ważnym krokiem. Wprawdzie umieścił on Słońce w centralnym miejscu we wszechświecie, a wszystkie planety łącznie z Ziemią zdegradował do roli satelitów, zachował jednak antyczny ruch planet po okręgu. Jego heliocentryczny system świata zakładał nawet dwukrotnie większą liczbę takich okręgów niż geocentryczny system Ptolemeusza. Dopóki nie rozważamy fizycznej rzeczywistości ciał

niebieskich, krok ten jest mniej rewolucyjny, niż się wydaje, bo oznacza tylko zmianę perspektywy. Jak wiemy z przekazu Archimedesa, heliocentryczny obraz świata głosił już w antyku Arystarch z Samos (od około 310 do około 230 r. p.n.e.). Dopiero po czasie okazało się, że kopernikański obraz świata może mieć zastosowanie w praktyce (na przykład w orientacji na pełnym morzu), ponieważ pomaga przewidywać ruchy gwiazd. Obraz ten nie zostałby jednak nigdy w pełni zaakceptowany, gdyby nie wiązało się z nim dążenie do prawdy. W tym celu konieczne było zastosowanie fizyki w odniesieniu do ciał niebieskich.

Przyczynił się do tego Galileusz (1564–1642), który w 1609 roku skierował swoją lunetę w niebo i w dziele *Sidereus Nuncius* (*Gwiezdny posłaniec* – wg przekładu ks. Krzysztofa Czapli w „Semina Scientiarum" 2005, nr 4) opisał to, co dostrzegł. Galileusz oraz poinformowany o tym wydarzeniu Johannes Kepler (1571–1630) uznali, że poza widocznym gołym okiem światem istnieje inny świat z nieznanymi do tej pory ciałami niebieskimi, takimi jak choćby księżyce Jowisza. Oznaczało to, że świat jest nieskończony, a przez to wiedza o nim może być niewyczerpana. Konstruowanie coraz lepszych lunet spowodowałoby odkrycie nowych gwiazd i zjawisk. Poznanie już nigdy więcej nie mogło być traktowane jako skończone, a w badaniach naukowych pojawił się czynnik czasu.

Użycie lunety i wkład w kopernikańską wizję świata to jedynie część dokonań Galileusza. Innym, nie mniej ważnym dla rozwoju nauki, było wprowadzenie pojęcia abstrakcji. Opublikowane w 1638 roku główne dzieło Galileusza *Discorsi*, traktujące o mechanice (szczegółowy tytuł wg przekładu Feliksa Kucharzewskiego, 1930 – *Rozmowy i dowodzenia matematyczne w zakresie dwóch nowych umiejętności dotyczących mechaniki i ruchów miejscowych*), jest nacechowane, jak to określał Blumenberg, siłą rozumu skierowaną przeciwko oglądowi. Inaczej niż u Arystotelesa, nie chodziło o możliwie dokładny opis świata widzianego na co dzień, lecz o konkretne eksperymenty przy dobrze zdefiniowanych założeniach. W poszczególnych przypadkach nie muszą to być eksperymenty przeprowadzone w rzeczywistości, ale po prostu

eksperymenty myślowe, co u Galileusza nie było wcale rzadką praktyką. Od tego momentu na znaczeniu zyskała sytuacja *wyidealizowana*. Galileusz był w stanie sformułować swoje prawa swobodnego spadania, pomijając zewnętrzny opór powietrza, co dla Arystotelesa było nie do pomyślenia. Taka wyidealizowana sytuacja musiała zostać opisana językiem matematyki. Bez tych dwóch innowacji, eksperymentów i matematyki, dalszy rozwój nauki nie byłby możliwy.

Galileusz nie mógł jednak całkowicie uwolnić się od spuścizny Arystotelesa, nadal trzymając się kurczowo przekonania o kołowych orbitach ciał niebieskich. Skończył z tym dopiero Kepler. W wyniku żmudnej, drobiazgowej pracy udało mu się z ogromu danych zebranych przez duńskiego astronoma Tychona Brahego (1546–1601) w czasach poprzedzających wynalezienie teleskopu wyodrębnić swoje tak zwane prawa Keplera dotyczące ruchu planet. Pierwsze z tych praw mówi, że planety poruszają się wokół Słońca nie po orbitach kołowych, lecz eliptycznych. Kepler uzyskał tę pewność dzięki obserwacji ruchu Marsa. Stwierdził on niewielkie odchylenie wynoszące około ośmiu minut kątowych (odpowiada to około jednej czwartej średnicy tarczy Księżyca w pełni) w porównaniu z ruchem po orbicie. Dzięki wprowadzeniu elipsy Kepler rozprawił się ze starym obrazem świata w sposób o wiele bardziej radykalny niż w przypadku odkryć Kopernika. Słońce, którego pozycja wyróżnia się tym, że znajduje się w centrum wszystkich elips, stało się nie tylko centralnym ciałem niebieskim, lecz zaczęło być również uznawane za jedną z możliwych przyczyn ruchu planet. Pomimo wagi odkrycia Keplera początkowo nie cieszyło się ono zbyt dużym zainteresowaniem.

W hermetycznym systemie praw natury to Isaacowi Newtonowi (1643–1727) przypadło w udziale udowodnienie ruchu eliptycznego. Jednocześnie okazało się, że inne krzywe stożkowe (parabole i hiperbole) są rozwiązaniami podstawowych równań. Za przykład może posłużyć wiele komet, które poruszają się po takich orbitach. Jedną z nich jest choćby kometa McNaughta, którą mogliśmy podziwiać na południowym niebie w styczniu 2007 roku. Parabole i hiperbole są jednak nieskończenie

długimi krzywymi, przez co nieskończoność nieuchronnie wdarła się do opisu natury. Newton jako pierwszy przedstawił jednorodny opis natury, składający się z uniwersalnych praw leżących w równym stopniu u podstaw świata ziemskiego, *jak i* niebiańskiego. Ta sama siła ciężkości sprawia, że jabłko spada z drzewa, jak i że Księżyc utrzymuje się na swojej orbicie wokół Ziemi. Według Einsteina absolutnie wyjątkowo musiało smakować odkrycie, że przyczyną ruchów ciał niebieskich jest dobrze znana nam siła ciężkości.

Newton opublikował swoje fundamentalne zasady dynamiki w 1687 roku w dziele pod tytułem *Philosophiae naturalis principia mathematica* (*Matematyczne zasady filozofii naturalnej* – wg przekładu dr. Sławomira Brzezowskiego, 2015), zwanym krótko *Principia*. Najważniejszą myślą tam zawartą jest stwierdzenie, że to nie ruch jednostajny prostoliniowy wymaga wyjaśnień, lecz jego odchylenie. Za takie odchylenie odpowiedzialne są *siły*, które niezależnie od zasad należy wyjaśnić. Najsłynniejszym przykładem jest właśnie siła grawitacji, która odpowiada za spadanie kamienia i ruch ciał niebieskich. Pierwsza zasada Newtona (jego *Lex Prima*) mówi, że ciało pozostaje w spoczynku lub porusza się ruchem jednostajnym prostoliniowym, czyli ze stałą prędkością, gdy nie działają na nie żadne siły. Także w tej *zasadzie bezwładności* pojawia się pojęcie nieskończoności, gdyż naturalne są już nie ruchy obrotowe, lecz prostoliniowe i jednostajne, jakby biegnące w nieskończoność. Chodzi tutaj przede wszystkim o sytuację wyidealizowaną, ponieważ nie są brane pod uwagę inne ciała obecne we wszechświecie ani ich siły grawitacji.

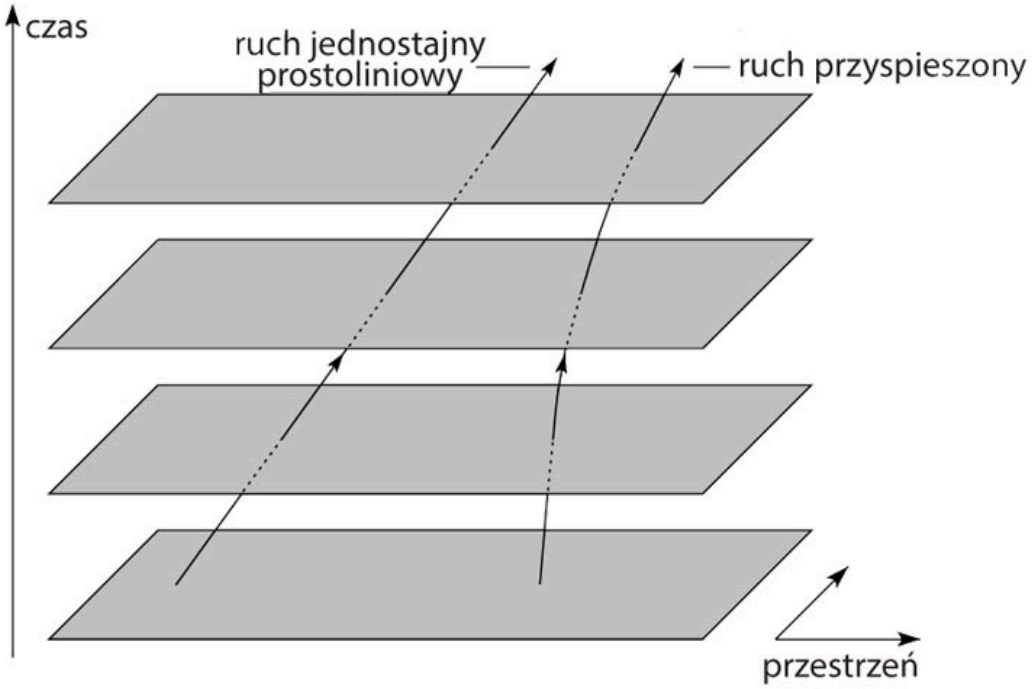
Zasada bezwładności obowiązuje oczywiście w określonych układach odniesienia, przy czym „układ odniesienia" ma dotyczyć stanu ruchu obserwatora. Gdy obserwator przyspiesza w stosunku do układu, w którym obowiązuje zasada bezwładności, to ciało obserwowane nie porusza się już ruchem jednostajnym z punktu widzenia przyspieszającego obserwatora. To nadzwyczajna cecha natury, że w ogóle istnieją takie doskonałe układy odniesienia, w których *wszystkie* wolne od oddziałujących sił ciała poruszają się ruchem jednostajnym prostoliniowym. Owe szczególne

układy nazywane są *układami inercjalnymi*, od łacińskiego słowa „inertia” oznaczającego bezwładność.

Druga zasada Newtona (jego *Lex Secunda*) pokazuje, *jak* należy opisać odchylenie od ruchu bezwładnego. Zasada ta uwzględnia zasadnicze pojęcie *siły*. Wpływ otoczenia zewnętrznego na ciało jest opisany przez jedną wielkość zwaną siłą. Jest ona odpowiedzialna za odchylenie od ruchu bezwładnego, a wynika z wpływu świata zewnętrznego. Za reakcję ciała na siłę odpowiada charakterystyczna dla ciała wielkość zwana *masą bezwładną*. Jest ona miarą ilości materii, z której składa się ciało. Jeżeli ciało przestaje poruszać się ruchem jednostajnym prostoliniowym, następuje jego przyspieszenie. Słynna zasada Newtona głosi wtedy, że masa razy przyspieszenie jest równe oddziałującej sile. Siła może powstawać w wyniku zderzenia bądź działać natychmiastowo na dowolną odległość. Doskonałym przykładem takiej siły jest siła grawitacji. Zgodnie z prawem powszechnego ciążenia Newtona dwa ciała we wszechświecie przyciągają się z siłą odwrotnie proporcjonalną do kwadratu odległości, czyli wraz z rosnącą odległością siła wyraźnie się zmniejsza. Jej wielkość określana jest przez stałą grawitacyjną (oznaczaną literą *G*), którą ustalił dopiero Henry Cavendish w 1798 roku. W tym celu przeprowadził on w Cambridge eksperyment laboratoryjny, podczas którego dwie masy próbne przymocowane do skrętnej nici były przyciągane przez masy grawitacyjne. Na podstawie otrzymanego po raz pierwszy wyniku oszacowano masę Ziemi i Słońca.

Jeżeli w przypadku tej siły chodzi o siłę grawitacji, pochodzącą od innego ciała, to jest ona proporcjonalna do pewnej wielkości danego ciała, którą nazywa się *masą grawitacyjną*. Zdumiewające jest to, że obie te masy, bezwładna i grawitacyjna, są równej wielkości, choć w formalizmie mechaniki Newtona nic na to nie wskazuje. Równość tę stwierdza się w wyniku obserwacji. Fakt ten był już znany Galileuszowi, jednak dopiero Einstein wyniósł zasadę równoważności masy bezwładnej i grawitacyjnej do rangi zasady fundamentalnej i oparł na niej swoją ogólną teorię względności, będącą pierwszym rozwinięciem mechaniki Newtona.

Do sformułowania swoich zasad Newton potrzebował pojęć przestrzeni absolutnej i czasu absolutnego. Według Einsteina Newton dostrzegł, że zarówno wielkości geometryczne (odległości między punktami materialnymi), które można zaobserwować, jak i ich przebieg w czasie nie oddają w pełni charakterystyki fizycznej ruchu. Szczególnie widoczne stało się to w słynnym eksperymencie Newtona z wiadrem. Wiadro wypełnione wodą wprowadza się w ruch obrotowy. Można zauważyć, że początkowo woda (przy maksymalnym ruchu względnym między ścianką wiadra a powierzchnią wody) pozostaje w spoczynku, potem jednak (podczas wspólnej rotacji wody i ścianki wiadra) unosi się przy ściankach wiadra. Co jest przyczyną takiego unoszenia się powierzchni wody? Z uwagi na to, że ruch względny już nie występował, Newton mógł sobie wyobrazić rotację wyłącznie absolutnie, tzn. w odniesieniu do przestrzeni absolutnej. Analogicznie jego *Lex Secunda* obowiązuje tylko wtedy, gdy pojawiający się w niej parametr czasu, na podstawie którego definiowane jest przyspieszenie, przyjmuje formę doskonałą. Według Newtona odzwierciedlało to istnienie czasu absolutnego. Rysunek 1.1 przedstawia newtonowskie pojęcie przestrzeni i czasu: następujące po sobie jednostajnie kopie trójwymiarowej przestrzeni absolutnej w czasie absolutnym. Przestrzeń jest płaska; obowiązują zasady geometrii euklidesowej. Kolejność ta jest obiektywna, gdyż w przypadku każdej pary zjawisk można obiektywnie rozstrzygnąć, czy występują one równocześnie, czy nie. Świat zewnętrzny rozwija się *w* czasie; równoczesność jest pojęciem absolutnym.



1.1. Newtonowskie pojmowanie przestrzeni i czasu. Czasoprzestrzeń można określić jako rozwój przestrzeni absolutnej (na rysunku jeden wymiar jest ukryty) w czasie absolutnym.

Na rysunku zaznaczono także kilka *linii świata*. Są to krzywe, które powstają, gdy śledzimy punkty przestrzenne w czasie. Proste, odpowiadające ruchowi bezwładnemu, są określone przez *Lex Prima*. Zakrzywiona linia świata odpowiada natomiast ruchowi przyspieszonemu. Zakłada się przy tym, że rozszerzalność ciał na ogół nie musi być uwzględniana, dlatego też często mówi się o „punktach materialnych" lub „cząstkach punktowych". Dzięki temu ciału można przyporządkować precyzyjnie wyznaczone miejsce i tym samym dobrze zdefiniowaną orbitę. Przyjmuje się, że rozszerzone ciało składa się z wielu punktów materialnych.

Układy inercjalne są układami doskonałymi. Jak jednak w praktyce odnaleźć układ inercjalny? Na to pytanie odpowiedział w 1885 roku Ludwig Gustav Lange (1863–1936), nazywany przez Maxa von Lauego „niesłusznie zapomnianym". Lange, od którego wywodzi się pojęcie

układu inercjalnego, obserwował trzy ciała wyrzucane równocześnie z jednego punktu w trzech różnych kierunkach, przy czym kierunki te nie powinny leżeć w jednej płaszczyźnie. Pod warunkiem braku oddziaływania jakichkolwiek sił (czyli przy domyślnym założeniu, że wszystkie istniejące siły zanikają dostatecznie szybko wraz z odległością) Lange pokazał, że można dzięki temu zdefiniować układ inercjalny. Wszystkie inne ciała, na które nie oddziałują żadne siły, poruszają się w stosunku do tak określonego układu ruchem jednostajnym prostoliniowym. W ten oto sposób Lange mógł dokładnie oddzielić prawa natury od przyjętych konwencji. Brytyjski fizyk i matematyk Peter Guthrie Tait (1831–1901) zaproponował alternatywne rozwiązanie. Pokazał on, że potrzebne są co najmniej cztery konfiguracje co najmniej trzech poruszających się punktów materialnych, na które nie oddziałują żadne siły, by zrekonstruować układ inercjalny. Przy zastosowaniu tej metody udaje się także rekonstrukcja przynależnej do niego inercjalnej skali czasu, czyli newtonowskiego czasu absolutnego. Należy jednak przy tym założyć, że newtonowskie równania ruchu obowiązują w odniesieniu do *tej* skali czasu. Rekonstrukcja czasu absolutnego należała przez wieki do zadań astronomii. Na podstawie równań obliczano położenie ciał niebieskich, jakie przyjmą one w różnym czasie w przyszłości. Z faktycznych obserwacji tych położeń można było dosłownie odczytać czas. Tak określony czas nazywamy *czasem efemeryd*. Nazwa pochodzi od efemeryd, czyli tabel wyszczególniających położenia ciał niebieskich. W pewnym sensie można zatem powiedzieć, że czas tkwi w danej chwili. Powrócimy do tego stwierdzenia przy omawianiu grawitacji kwantowej (część 5).

Równania Newtona w imponujący sposób opisywały wszelkie zjawiska mechaniki łącznie z mechaniką nieba, które odkrywano w ciągu dwustu lat po opublikowaniu *Principia*. Ponowne odnalezienie odkrytej w 1801 roku, lecz szybko zgubionej planety karłowatej Ceres, dzięki obliczeniom przeprowadzonym przez Gaussa, a także odkrycie planety Neptun to zaledwie dwa z długiej listy sukcesów. Nie byłyby one jednak możliwe bez matematycznej przystępności zasad Newtona, który samodzielnie

(równocześnie z Leibnizem, lecz niezależnie od niego) opracował równania potrzebne do ich matematycznego zapisu. W matematyce równania te nazywa się „równaniami różniczkowymi". Opisują one zmiany wielkości sąsiadujących punktów w czasie. Określenie przebiegu ruchu jest jednak możliwe dopiero po wybraniu dowolnie zadanych warunków początkowych, co także należało do szczególnych właściwości nowego obrazu świata. Owe warunki początkowe dotyczą położeń i prędkości obserwowanych ciał w wybranym czasie. Noblista Eugene Paul Wigner (1902–1995) uznał rozdzielenie opisu natury na właściwe prawa nią rządzące i dowolnie zadane warunki początkowe za największe odkrycie Newtona, ponieważ pokazuje ono, że do określenia przebiegu zdarzeń w przyrodzie poza samymi prawami niezbędne są również wspomniane warunki początkowe.

Newton był świadomy ograniczeń stworzonych przez siebie pojęć, czyli przestrzeni absolutnej i siły działającej na odległość. Swoimi przemyśleniami podzielił się z filologiem klasycznym Richardem Bentleyem: „Przypuszczenie, że grawitacja jest przyrodzoną, nieodłączną i istotną cechą materii sprawiającą, iż jedno ciało może działać na drugie na odległość, bez pośrednictwa czegoś, co przekazywałoby między nimi działanie i siłę, jest dla mnie tak wielkim absurdem, że nie sądzę, by ktokolwiek obdarzony zdolnością rozważania zagadnień filozofii mógł to kiedykolwiek przyjąć" (wg przekładu Andrzeja Kajetana Wróblewskiego, 2006). Dla Newtona było jednak oczywiste, że bez tych pojęć nigdy nie opracowałby swojej mechaniki. Tym samym krytyka tych pojęć przez jemu współczesnych, choć uzasadniona, trafiała w próżnię. Gottfried Wilhelm Leibniz (1646–1716) traktował przestrzeń i czas jedynie jako zasady porządku, a nie jako wielkości absolutne. Nie mógł jednak zaproponować żadnej przekonującej alternatywy dla mechaniki Newtona. Dopiero ogólna teoria względności Einsteina, omówiona w kolejnym rozdziale, ostatecznie zerwała z pojęciem przestrzeni absolutnej Newtona.

Od przestrzeni i czasu do czasoprzestrzeni: szczególna teoria względności Einsteina

W swoim *Dialogu o dwu najważniejszych układach świata: ptolemeuszowym i kopernikowym* z 1632 roku (wg przekładu Edwarda Ligockiego, 1962) Galileusz opisuje interesujący eksperyment myślowy. Parę osób ma zgromadzić się w zamkniętym pomieszczeniu pod pokładem nieporuszającego się statku i przeprowadzić kilka obserwacji oraz eksperymentów mechanicznych (np. eksperymenty z rzutami). Potem, jak mówi Galileusz głosem swojego alter ego Salviatiego: „(...) pozwólcie popłynąć statkowi z dowolną prędkością. Nie dostrzeżecie u wszystkich wymienionych zjawisk żadnej zmiany, o ile ruch będzie jednostajny i pozbawiony kołysania. Z żadnego z nich nie wywnioskujecie, czy statek stoi, czy płynie". W ten sposób Galileusz opisał *zasadę względności* mechaniki. W skrócie zasada ta postuluje równoważność wszystkich układów inercjalnych, to znaczy niemożliwość wyodrębnienia określonego układu inercjalnego na podstawie przeprowadzonych pomiarów mechanicznych.

Pytanie, czy zasada względności obowiązuje także w odniesieniu do zjawisk niemechanicznych, naprowadziło Alberta Einsteina (1879–1955) w 1905 roku na jego *szczególną teorię względności*. Kluczem do odpowiedzi na to pytanie były fale elektromagnetyczne. Zalicza się do nich w szczególności rozchodzenie się światła. Na początku XIX wieku dołączono do tego obraz fal, choć wówczas traktowano światło jako byt odrębny od zjawisk elektrycznych i magnetycznych. W tym czasie wyobrażano sobie, że do rozchodzenia się fal może dochodzić wyłącznie za pośrednictwem jakiegoś drgającego nośnika. W falach wodnych drgają pojedyncze cząsteczki wody, a w przypadku fal dźwiękowych w powietrzu poszczególne cząsteczki powietrza. Co zatem drga podczas rozchodzenia się światła? Z uwagi na to, że nie stwierdzono obecności żadnego oczywistego nośnika, opierając się przy doborze słów na Arystotelesie, postulowano istnienie hipotetycznego *eteru*, który jest rozpowszechniony we wszechświecie w równomiernym stopniu i przenika każdą zwykłą

materię. Z obserwacji duńskiego astronoma Olego Rømera (1644–1710) z lat 1672–1676 wiedziano już wtedy, że światło charakteryzuje się wprawdzie skończoną, ale bardzo wysoką prędkością. W 1672 roku Rømer stwierdził bowiem, że odległość Ziemi od Słońca (dziś określana mianem jednostki astronomicznej) można uznać za wystarczającą miarę do określenia prawdopodobnie skończonej prędkości światła. Rømer obliczył prędkość światła na podstawie przesunięć w czasie, jakie zaobserwował podczas zaćmień księżyców Jowisza. Opóźnienia te wynikają ze zróżnicowanego czasu potrzebnego światłu na pokonanie dystansu z Jowisza do Ziemi, gdy Ziemia znajduje się w różnych punktach na swojej orbicie wokół Słońca. Dzisiejsza wartość prędkości światła w próżni wynosi prawie 300 000 kilometrów na sekundę – wartość tę wyraża się w skrócie za pomocą litery *c*. Przy tak dużej prędkości eter musiałby wykazywać niezwykłą wytrzymałość. Poza tym jak miałby być mimo to niewidoczny?

Jeśli ktoś porusza się z określoną prędkością względem eteru, powinien zarejestrować tym samym zmienioną prędkość światła, zmniejszoną lub powiększoną o prędkość własną. Z uwagi na to, że Ziemia na swojej rocznej orbicie wokół Słońca zmienia w szczególności kierunek swej prędkości, zjawisko to powinno wyrażać się w rocznych wahaniach mierzonej prędkości światła. Oczywiście można sobie także wyobrazić, że Ziemia wraz ze swoją atmosferą ciągnie za sobą eter i dlatego nie można zmierzyć zróżnicowanej prędkości światła. Przeciwko tej hipotezie świadczyła jednak aberracja światła – fenomen znany już od 1728 roku. Pod tym pojęciem rozumiane jest przesunięcie pozycji gwiazd, spowodowane rocznymi zmianami prędkości Ziemi (nie zmianami jej położenia). Jeżeli zatem ruch światła przebiegał w eterze, Ziemia nie mogła ciągnąć ze sobą eteru, a już z pewnością nie w całości. Żadna z prób potwierdzenia ruchu względem eteru nie powiodła się.

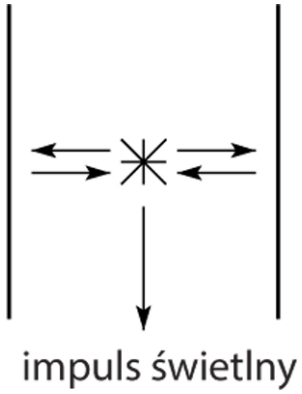
Problem stał się bardziej palący, gdy udało się opracować jednolitą teorię dla wszystkich zjawisk elektrycznych i magnetycznych, czyli elektrodynamikę. Jej podstawowe równania zostały opublikowane w 1873

roku przez Jamesa Clerka Maxwella (1831–1879) w słynnym dziele na temat elektryczności i magnetyzmu *A Treatise on Electricity and Magnetism*. Jedną z najważniejszych prognoz równań Maxwella było istnienie fal elektromagnetycznych, które rozprzestrzeniają się z prędkością światła *c*. Szczególnym przypadkiem fali elektromagnetycznej okazało się światło. Bezpośredni dowód na istnienie fal elektromagnetycznych przyniosły epokowe eksperymenty Heinricha Hertza (1857–1894) z 1887 roku, które pomogły teorii Maxwella w dokonaniu przełomu. Stały się także początkiem bezprzewodowego przekazywania informacji – już w 1901 roku można było przesyłać fale radiowe przez Atlantyk.

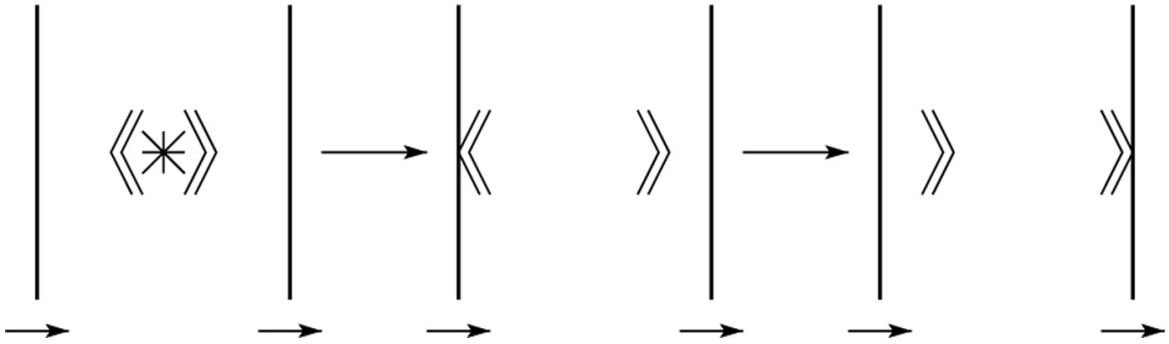
Równania Maxwella zdawały się obowiązywać jednak tylko w jednym określonym układzie odniesienia – w „układzie eterycznym", w którym światło rozprzestrzenia się z prędkością *c*. Czy zatem zasada względności Galileusza nie obowiązuje w przypadku zjawisk elektrodynamicznych? Chociaż nieświadomie, to pytanie zadawał sobie już szesnastoletni Einstein w szkole kantonalnej w Aarau. Później pisał o tym tak: „Podczas tego roku spędzonego w Aarau przyszło mi do głowy pytanie: gdyby tak pobiec za falą świetlną z prędkością światła, to miałoby się przed sobą niezależne od czasu pole falowe. Jednak wygląda na to, że coś takiego nie istnieje! Był to pierwszy dziecięcy eksperyment myślowy, mający związek ze szczególną teorią względności. Odkrywanie nie jest dziełem logicznego myślenia, choć produkt końcowy jest powiązany z formą logiczną". Takie niezależne od czasu pole falowe nie jest wynikiem równania Maxwella – podejrzewano naruszenie zasady względności.

Szczególna teoria względności Einsteina z 1905 roku powstała z próby połączenia ze sobą zasady względności z elektrodynamiką. Einstein uważał, że wszystkie zasady fizyki, nie tylko mechaniczne, powinny mieć tę samą formę we wszystkich układach inercjalnych. Układ inercjalny można uzyskać za pomocą wyżej opisanej metody Langego lub przez odniesienie do układu pozagalaktycznych fal radiowych, czyli *International Celestial Reference System*.

Jak zasada względności ma się jednak do prędkości światła *c* w równaniach Maxwella? Czy właśnie podczas przejścia do innego układu inercjalnego prędkość światła nie ulegnie zmianie? Z uwagi na to, że wszystkie próby udowodnienia ruchu względem eteru spełzły na niczym, Einstein stwierdził po prostu, że prędkość światła we *wszystkich* układach inercjalnych równa się *c*. Sam eter stał się wielkością zbędną, którą bez wątpienia można było wyeliminować z fizyki. To, że postulat niezmiennej prędkości światła da się pogodzić z zasadą względności, Einstein wykazał poprzez dokładną analizę pojęcia czasu: w przypadku dwóch zdarzeń, do których dochodzi w różnych miejscach, nie można powiedzieć, że następują one w tym samym czasie. Najlepiej pokazać to na prostym przykładzie (rys. 1.2).



1.2 a). Statyczne lustra równoległe



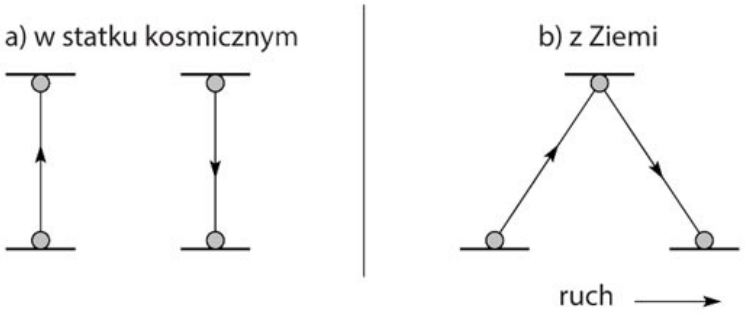
1.2 b). Równoległe lustra poruszające się ze stałą prędkością w prawą stronę

Widzimy dwa ustawione pionowo lustra, które są równoległe względem siebie i poruszają się w prawo ze stałą prędkością. W środku, między lustrami, znajduje się źródło światła, z którego równocześnie wysyłane są impulsy świetlne w przeciwnych kierunkach – w stronę luster. Przypatrzmy się najpierw obserwatorowi, który nie porusza się względem luster. Obserwator ten rejestruje równoczesne dotarcie impulsów światła do luster (rys. 1.2 a). Jeśli jednak spojrzeć na to z zewnątrz, widać, że jeden impuls przemieszcza się do najbliższego lustra i dociera do niego, *zanim* drugi impuls dotrze do drugiego lustra. Powodem tego jest prędkość światła, która zawsze wynosi c (rys. 1.2 b). Zewnętrzny obserwator nie dostrzeże zatem równoczesnego dotarcia impulsów światła do luster. Po odwróceniu sytuacji oba impulsy powracają równocześnie do środka. Równoczesność ta dotyczy tym razem obu obserwatorów – równoczesność w tym samym punkcie przestrzeni jest więc w dalszym ciągu absolutna.

W przeciwieństwie do mechaniki Newtona nie istnieje zatem równoczesność absolutna! Po tym odkryciu siły Newtona działające na odległość zostały zapomniane. W przypadku oddziaływania elektromagnetycznego stało się oczywiste, że zakłócenia rozchodzą się w formie fal elektromagnetycznych z prędkością światła c. Do losów wyobrażenia Newtona o natychmiastowo oddziałującej sile ciężkości jeszcze wrócimy. W świecie fizyki pojawiło się pojęcie *pola*, które, jak się później okazało, stało się kluczowe dla dalszego rozwoju nauki. Początkowo mówiono jedynie o dwóch siłach: sile między dwiema masami lub sile między dwoma ładunkami elektrycznymi. Michael Faraday (1791–1867) wprowadził do opisu swoich eksperymentów pojęcie „linii pola siłowego” i tym samym uniezależnił je od sił oddziałujących na ładunki. Pole elektryczne i pole magnetyczne posiadają własną rzeczywistość. Maxwell i Faraday mogli je sobie wyobrazić wyłącznie w odniesieniu do materialnej substancji nośnej, czyli właśnie eteru. We współczesnym ujęciu podstawowe pola zdefiniowane są bezpośrednio w przestrzeni i czasie bez posiłkowania się nośnikiem. Odrzucenie takiego nośnika było jednym z największych i najtrudniejszych dokonań w rozwoju fizyki.

Rezygnacja z koncepcji równoczesności absolutnej ma zasadnicze znaczenie dla zależności między przestrzenią a czasem. Podczas przejścia z jednego układu inercjalnego do drugiego zmieniają się współrzędne przestrzeni *i* współrzędna czasu. W mechanice klasycznej czas pozostaje niezmienny. Konsekwencją tej mieszanki przestrzeni i czasu jest *dylatacja czasu*. Zegar poruszający się względem innego zegara porusza się wolniej względem niego. Najłatwiej będzie pokazać ten stan rzeczy za pomocą zegara świetlnego, z którym spotkaliśmy się już wcześniej w nieco zmienionej formie (zob. rys. 1.3).

Każdy zegar opiera się na okresowym przebiegu. W zegarze świetlnym promień świetlny odbijany jest na przemian między dwoma lustrami. Ten naprzemienny ruch odpowiada w pewnym sensie „tykaniu" zegara. Wyobraźmy sobie, że zegar znajduje się we wnętrzu statku kosmicznego, oddalającego się od Ziemi ze stałą prędkością. W przeciwieństwie do rysunku 1.2 ruch powinien przebiegać równolegle do luster. W układzie spoczynkowym statku kosmicznego promień świetlny porusza się prostopadle do luster. W tym celu potrzebny jest określony czas (rysunek z lewej). Jeśli całą tę sytuację obserwujemy z Ziemi, okazuje się, że promień świetlny biegnie ukośnie i musi przebyć dłuższą drogę (rysunek z prawej). Skoro prędkość światła jest zawsze taka sama, promień świetlny potrzebuje do przebycia dłuższej drogi także więcej czasu, a więc „zegar" porusza się wolniej. Analogicznie będzie oczywiście w przypadku zegara świetlnego umieszczonego na Ziemi i obserwowanego ze statku kosmicznego. Zjawisko to dotyczy wyłącznie ruchu *względnego*.



1.3. Zegar świetlny

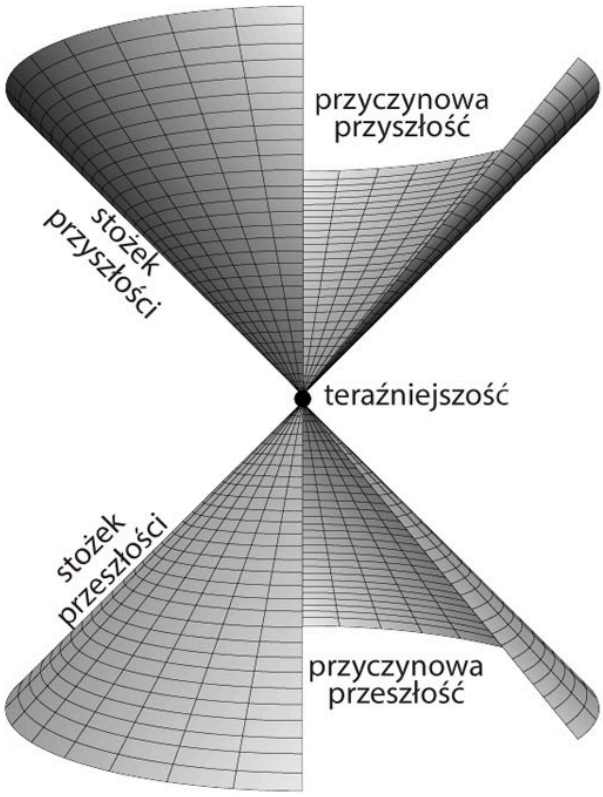
W związku z zasadą względności każdy poruszający się zegar musi spóźniać się o tę samą wartość. W przeciwnym wypadku poprzez porównanie różnych zegarów w ramach układu inercjalnego można byłoby stwierdzić, z jaką prędkością absolutną porusza się dany układ, co byłoby niezgodne z zasadą względności. Dylatacja czasu jest właściwością ruchu względnego w przestrzeni i czasie i nie zależy od specjalnych zegarów. Analogiczny efekt dotyczy także przestrzeni. Okazuje się, że linijka poruszająca się względem spoczywającej linijki o takiej samej długości jest krótsza. Te i inne zjawiska szczególnej teorii względności zostały dokładnie zbadane i potwierdzone dzięki wielu eksperymentom przeprowadzonym po 1905 roku. W przypadku niskich (w porównaniu z prędkością światła) prędkości twierdzenia teorii względności nie różnią się od twierdzeń mechaniki klasycznej. Z uwagi na to, że w historycznych eksperymentach i obserwacjach życia codziennego mamy do czynienia wyłącznie z niskimi prędkościami, staje się jasne, dlaczego zjawiska tego typu nie odgrywają tam żadnej roli.

Mianem wesołego i przekonującego określił Einstein zastosowanie zasady względności, które doprowadziło go do prawdopodobnie najsłynniejszego wzoru fizycznego, mówiącego, że energia i masa są równoważne, czyli $E = mc^2$. Współczynnikiem przeliczeniowym jest właśnie kwadrat prędkości światła, a zatem ogromny współczynnik, z powodu którego już małym masom odpowiadają potężne ilości energii. Wesołe i przekonujące zastosowanie zasady względności wyglądało

następująco: Einstein obserwował ciało, które w swoim układzie spoczynkowym wysyła w przeciwnych kierunkach dwie fale świetlne o równej częstotliwości. Jeśli przy odpowiednim zastosowaniu równań teorii względności sytuację tę ocenia się z perspektywy układu inercjalnego, poruszającego się względem tego ciała, można stwierdzić, że bezwładna masa i energia muszą zostać ustalone. W znanym przykładzie rozszczepienia jądra atomowego produkty rozszczepienia są trochę lżejsze od materiału wyjściowego, co wystarcza do wyprodukowania energii w elektrowni atomowej. Pozornie niewinna zasada względności może mieć więc bardzo poważne konsekwencje. W eksperymentach przeprowadzonych w 2005 roku między innymi w Massachusetts Institute of Technology w Stanach Zjednoczonych i w Instytucie Lauego-Langevina w Grenoble we Francji ważność słynnego wzoru Einsteina została potwierdzona przy niepewności pomiaru rzędu 0,00004 procenta. W ramach eksperymentu zmierzono utratę masy, do której dochodzi w jądrach pierwiastków krzemu i siarki podczas emitowania wysokoenergetycznego promieniowania gamma.

Bez słynnego wzoru Einsteina nie istniałoby życie na Ziemi. Niezbędna do życia energia słoneczna powstaje w drodze syntezy jądrowej, wskutek której masa przekształca się w energię. Ilość energii wyemitowanej przez Słońce na sekundę odpowiada masie 4,4 miliona ton, o którą Słońce staje się lżejsze.

Eleganckie matematyczne sformułowanie szczególnej teorii względności zaprezentował Hermann Minkowski (1864–1909), jeden z dawnych nauczycieli Einsteina na politechnice w Zurychu, w 1908 roku podczas głośnego wykładu w Kolonii. Połączył on trójwymiarową przestrzeń i jednowymiarowy czas w jeden czterowymiarowy geometryczny twór – *czasoprzestrzeń*. Na jego cześć twór ten nazywa się „przestrzenią Minkowskiego” (a dokładniej: „czasoprzestrzenią Minkowskiego”). Punkty w tej przestrzeni zwane są *zdarzeniami*. Zamiast newtonowskiej kolejności trójwymiarowych przestrzeni w czasie absolutnym (rys. 1.1) istnieje inna struktura absolutna – *stożek świetlny* (rys. 1.4).

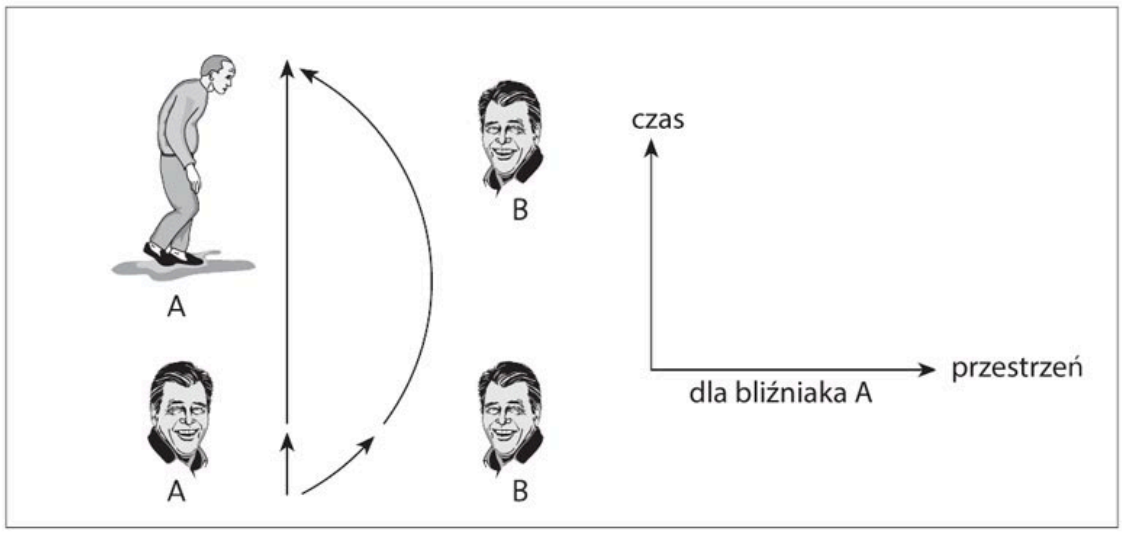


1.4. Stożek przyszłości i przeszłości zdarzenia

Stożek świetlny (a dokładniej: podwójny stożek) odnosi się do każdego zdarzenia. Charakteryzuje on wszystkie punkty w czasoprzestrzeni, do których mogą dotrzeć promienie świetlne począwszy od tego zdarzenia („stożek przyszłości") lub które mogą one objąć, powracając do tego zdarzenia („stożek przeszłości"). Z teorii względności wynika, że ciała materialne poruszają się zawsze z prędkością mniejszą od prędkości światła, ponieważ w przeciwnym razie potrzebowałyby nieskończonej ilości energii, by osiągnąć prędkość światła. Linie świata ciał materialnych przebiegają zawsze wewnątrz stożka świetlnego. Samo światło biegnie po powierzchni bocznej stożka. Przyczynowa struktura czasoprzestrzeni Minkowskiego nadal jest absolutna, a tylko przestrzeń i czas już takie nie są. Cały obszar poza stożkiem świetlnym nazywany jest teraźniejszością, bo dla każdego zdarzenia w tym obszarze można znaleźć układ inercjalny,

w stosunku do którego równocześnie następuje zdarzenie na początku stożka świetlnego. Jest to wyrazem utraty równoczesności absolutnej. „Czasoprzestrzeń” Newtona wynika z rysunku 1.4 i powstaje po otwarciu podwójnego stożka do środka, co odpowiadałoby istnieniu nieskończonej prędkości światła.

Szczególna teoria względności zajmuje się przede wszystkim ruchem jednostajnym prostoliniowym, ponieważ w tym przypadku obowiązuje zasada względności, a przejście między różnymi układami inercjalnymi jest proste do opisania. Teoria ta potrafi jednak również opisywać przyspieszenia, choć wiele osób przyprawia to o ból głowy. Można się o tym przekonać na przykładzie słynnego „paradoksu bliźniąt” (porównaj z rys. 1.5).



1.5. Paradoks bliźniąt

W przykładzie tym jeden z bliźniaków (na rysunku bliźniak B) opuszcza Ziemię w statku kosmicznym, podczas gdy jego brat (bliźniak A) pozostaje na Ziemi. Pasażer statku kosmicznego pędzi z ogromną prędkością w kosmos, a potem zawraca i pędzi z powrotem na Ziemię. Po powrocie stwierdza, że jego brat wygląda starzej od niego. Jak można to wyjaśnić? Z uwagi na to, że z powodu dylatacji czasu wszystkie zegary (a tym samym także wszystkie procesy fizjologiczne) w statku

kosmicznym pracowały wolniej niż na Ziemi, można oczekiwać, że astronauta w chwili swojego powrotu mniej się zestarzał. Ale czy nie działa to również w drugą stronę? Czy to właśnie nie ten bliźniak, który pozostał na Ziemi, porusza się względem brata tkwiącego w statku kosmicznym i to nie on będzie tym młodszym z nich dwóch? Haczyk w takim ciągu myślowym tkwi w tym, że astronauta doświadcza przyspieszenia, natomiast jego ziemski odpowiednik nie (w każdym razie w sporym przybliżeniu). By móc zawrócić, podróżujący bliźniak musi zmienić swoją prędkość. Zasada względności obowiązuje jednak tylko dla układów inercjalnych, a nie dla układów przyspieszonych. Z tego powodu symetria między bliźniakami jest zaburzona. Ze sformułowania Minkowskiego o czasoprzestrzeni łatwo odczytać, że ziemski bliźniak w chwili powrotu brata jest starszy. Upływający czas, który każdy z bliźniaków odczytuje na swoim zegarku (mowa tu o tak zwanym *czasie własnym*), można odczytać bezpośrednio z długości poszczególnych linii świata w przestrzeni Minkowskiego. Odwrotnie niż w klasycznej geometrii przestrzennej przyjmuje się jednak, że z największym ubytkiem czasu musi liczyć się obserwator znajdujący się na prostej linii świata. Każda krzywizna skraca czas własny. Z uwagi na to, że linia świata podróżującego bliźniaka jest zakrzywiona (zob. rys. 1.5), odpowiada jej krótszy czas własny w porównaniu z linią świata ziemskiego bliźniaka.

Nawet jeśli Einstein w odpowiedzi na matematyczne sformułowanie Minkowskiego miał stwierdzić, że nie rozumie już swojej teorii, odkąd zajmują się nią matematycy, to sam szybko podchwycił wyobrażenie czterowymiarowej czasoprzestrzeni. To właśnie ono okazało się kluczowe dla jego największego dokonania – ukończonej w 1915 roku ogólnej teorii względności i powiązanej z nią geometryzacji świata.

Geometryzacja świata:

ogólna teoria względności Einsteina

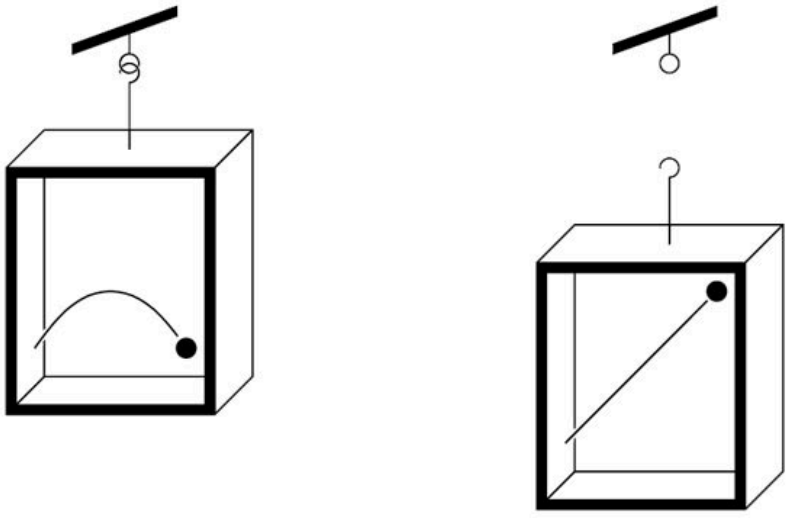
W *Discorsi* Galileusza Salviati wypowiada się na temat spadania. Twierdzi, że to niemożliwe, aby kamień o wadze 20 funtów miał spadać 10 razy szybciej niż ten o wadze dwóch funtów. Oba mogą spadać z wysokości 50 lub 100 łokci, a i tak spadną w tej samej chwili. Wypowiedź ta stała w sprzeczności z poglądami Arystotelesa, wyrażonymi w dialogu z Simplicio, w myśl których cięższe ciała spadają szybciej. Fakt, że wszystkie ciała spadają z taką samą prędkością, był już znany Galileuszowi. W rozdziale 1 nazwaliśmy ten stan rzeczy zasadą równoważności: masa grawitacyjna i masa bezwładna są tej samej wielkości. W związku z tym w równaniach ruchu obie masy zostają zniesione i ruch jest niezależny od masy. Newton potwierdził eksperymentalnie tę równoważność z dokładnością rzędu około jednego promila. Nie przypisywano jednak temu faktowi większego znaczenia. Według Einsteina Galileusz, Newton i późniejsi naukowcy wprawdzie przyjęli do wiadomości fakt równoważności masy bezwładnej i grawitacyjnej, ale go nie wykorzystali. Dla Einsteina stała się ona jednak punktem wyjścia do ogólnej teorii względności – współczesnej teorii grawitacji.

Zasada równoważności wskazała drogę do geometrycznego opisu czasoprzestrzeni. Einstein głosił tę zasadę w następującym uogólnionym kształcie: swobodnie spadający układ odniesienia jest względem wszystkich zjawisk w przyrodzie równoważny z układem odniesienia znajdującym się w przestrzeni, w której nie oddziałuje siła ciężkości. Z tego wynika oczywiście równość masy bezwładnej i grawitacyjnej, ale ponadto też pełna równoważność swobodnego spadania i nieważkości w odniesieniu do wszystkich możliwych eksperymentów. Należy jednak zaznaczyć, że równoważność ta obowiązuje tylko lokalnie, czyli tylko w przypadku układów odniesienia, które nie mają dużych rozmiarów. W ciągu ostatnich stu lat zasada równoważności została wielokrotnie

potwierdzona – niepewność pomiaru równości masy grawitacyjnej i bezwładnej wynosi mniej niż 10^{-12}. W ostatnich latach, również dzięki eksperymentom przeprowadzonym w kosmosie, dokładność ta jeszcze wzrosła.

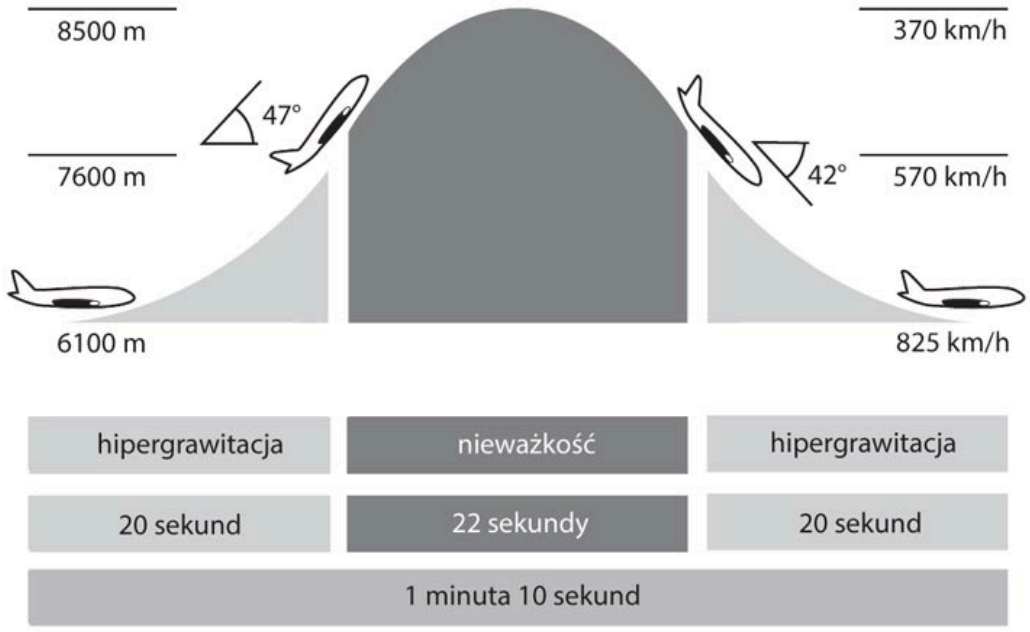
Dlaczego zasada równoważności prowadzi do geometryzacji świata? Dzieje się tak, ponieważ swobodnie spadający układ (ulubionym przykładem Einsteina była swobodnie spadająca winda) zachowuje się jak układ nieważki – powstaje lokalny układ inercjalny. Wolne od oddziałujących sił obiekty poruszają się w nim ruchem jednostajnym prostoliniowym (zob. rys. 1.6).

Inaczej niż w czasach Einsteina urzeczywistnienie swobodnego spadku z udziałem ludzi nie jest już obecnie niczym niezwykłym. Nie dotyczy to jedynie astronautów. Dzięki zwykłym samolotom można zrealizować lot paraboliczny, w którym samolot, jak w przypadku rzutu ukośnego, porusza się sam pod wpływem grawitacji, czyli swobodnie spada. (Parabola jest efektem obecności dodatkowej prędkości początkowej). Na rysunku 1.7 przedstawiono typowe parametry lotu parabolicznego, podczas którego można przeprowadzić fizyczne eksperymenty.



1.6. Zasada równoważności: swobodnie spadająca winda (po prawej) odpowiada windzie w stanie nieważkości. Winda zawieszona w polu ciężkości (po lewej) odpowiada windzie znajdującej się w ruchu przyspieszonym.

Z uwagi na to, że w układzie swobodnego spadania nie odczuwa się już grawitacji, ponownie obowiązują zasady szczególnej teorii względności. Dzieje się tak jednak tylko lokalnie! Gdybyśmy ponad Ziemią obserwowali takie swobodnie spadające windy, znacznie od siebie oddalone, stwierdzilibyśmy, że wszystkie przemieszczają się w kierunku punktu środka Ziemi. Nawet jeśli wystartują równolegle, to i tak zbliżą się do siebie i będą poruszać się względem siebie ruchem przyspieszonym. Z tego wynika, że wiele lokalnych układów inercjalnych już do siebie nie pasuje i nie może stworzyć globalnego, obejmującego wszystko układu inercjalnego. Globalny układ inercjalny odpowiada płaskiej czasoprzestrzeni, ponieważ obowiązują w nim zasady geometrii euklidesowej. Gdy wiele małych płaskich czasoprzestrzeni już do siebie nie pasuje, znaczy to, że cała czasoprzestrzeń jest zakrzywiona.



1.7. Lot paraboliczny: przez około 20 sekund samolot praktycznie swobodnie spada

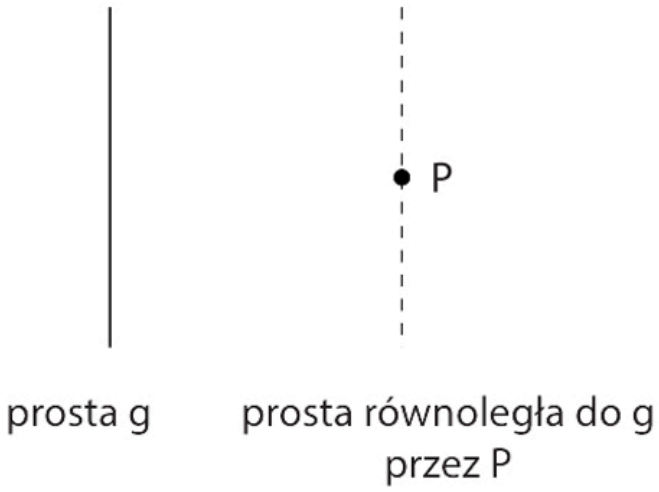
W tym miejscu pomoże nam analogia z obserwacji zwykłej przestrzeni. Na przykład powierzchnia ziemniaka jest z pewnością zakrzywiona, jednak można ją do siebie upodobnić dzięki wielu małym płaskim kawałkom, które można sobie wyobrazić jako małe znaczki pocztowe przyklejone do ziemniaka. Każdy znaczek jest płaski, choć cała powierzchnia jest zakrzywiona. Podobnie jest z czasoprzestrzenią, gdzie lokalne układy inercjalne przejmują rolę znaczków pocztowych.

W swobodnie spadającym układzie ciała poruszają się ruchem jednostajnym prostoliniowym, czyli po odcinkach prostych (zob. rys. 1.6). Z uwagi na to, że lokalne, swobodnie spadające układy nie pasują do siebie, również lokalne odcinki prostych nie mogą połączyć się w globalną prostą. Wynikiem tego połączenia jest jednak możliwie prosta krzywa. Zgodnie z pojęciem funkcjonującym w dziedzinie pomiarów geodezyjnych krzywe takie nazywane są *geodezyjnymi*. Są one uogólnieniami prostych w zakrzywionych przestrzeniach czy czasoprzestrzeniach. W przestrzeni są

to takie krzywe, które mają minimalną odległość między dwoma danymi punktami. Jeśli w kontekście zakrzywionych przestrzeni mówimy o prostych, to zawsze są to linie geodezyjne. W czasoprzestrzeni krzywe te mają największą możliwą odległość w odniesieniu do czasu własnego między dwoma zdarzeniami. Na linii świata, którą tworzy geodezyjna, mamy do czynienia z najdłuższym czasem własnym. Z tym ostatnim spotkaliśmy się już przy okazji omawiania paradoksu bliźniąt. Bliźniak pozostający na Ziemi porusza się (choćby w przybliżeniu) po geodezyjnej w czasoprzestrzeni Minkowskiego, podczas gdy astronauta porusza się ruchem przyspieszonym i dlatego leci wzdłuż krzywej, która nie jest geodezyjną i na której czas własny płynie wolniej.

Jak jednak zapisać matematycznie zakrzywioną przestrzeń? To pytanie zajmowało Einsteina od 1907 roku, gdy odkrył, że zasada równoważności może być kluczem do zrozumienia grawitacji, aż do opracowania podstawowych równań ogólnej teorii względności w 1915 roku.

Od początków matematyki w antycznej Grecji aż do XIX wieku dominującą pozycję zajmowała geometria euklidesowa. Składa się ona z szeregu aksjomatów, czyli bezpośrednio zrozumiałych postulatów, oraz wyciągniętych z nich wniosków, czyli właściwych matematycznych twierdzeń. We współczesnej matematyce rezygnuje się jednak z tej bezpośredniej zrozumiałości. Aksjomaty są formułowane w mniejszym lub większym stopniu dowolnie, w pewnym sensie jak w grze językowej. W geometrii euklidesowej szczególną rolę odgrywa słynny piąty aksjomat, który mówi, że w płaszczyźnie dla każdej prostej i dla każdego punktu poza tymi prostymi istnieje dokładnie jedna prosta równoległa do tej pierwszej, czyli jej nie przecinająca (zob. rys. 1.8).



1.8. W geometrii euklidesowej dla każdej prostej i dla każdego leżącego poza nią punktu P istnieje dokładnie jedna prosta przebiegająca przez P i równoległa do pierwszej prostej

Na przestrzeni wieków matematycy stale zadawali sobie pytanie, czy ten aksjomat jest niezależny, czy też można go wyprowadzić z innych aksjomatów. Wszystkie próby wyprowadzenia go spełzły jednak na niczym. Jego autentyczną niezależność udało się potwierdzić dopiero w XIX wieku poprzez jasne opracowanie geometrii spełniającej wszystkie aksjomaty Euklidesa oprócz piątego. Geometrie, w których nie obowiązuje piąty aksjomat Euklidesa, określa się mianem *nieeuklidesowych*. Pierwszych przykładów dostarczyli Nikołaj Łobaczewski (1792–1856), János Bolyai (1802–1860) i Carl Friedrich Gauss (1777–1855). Początkowo chodziło o geometrie, w których dla danej prostej istnieje nieskończenie wiele równoległych prostych przechodzących przez punkt leżący poza daną prostą. Taką geometrię nazywa się hiperboliczną, a jej przykładem jest powierzchnia siodła. Wyróżnia się ona ponadto tym, że suma kątów trójkąta jest mniejsza niż 180°. Są jeszcze geometrie sferyczne, w których nie istnieje równoległa dla danej prostej, a suma kątów trójkąta jest większa niż 180°. Sferycznych geometrii początkowo nie brano pod uwagę podczas rozpatrywania powyższego problemu,

a przyczyna takiego stanu rzeczy mogła tkwić w tym, że istnieją w nich punkty, które można połączyć za pomocą nieskończenie wielu sferycznych prostych (chodzi o diametralnie przeciwstawne punkty), to zaś narusza kolejny aksjomat Euklidesa o dwóch punktach, przez które przechodzi dokładnie jedna prosta.

Oczywiście zakrzywione powierzchnie, jak choćby powierzchnia kuli, były znane już od czasów antyku, ale można było wyobrazić je sobie jedynie poprzez ich osadzenie w przestrzeni trójwymiarowej, w której obowiązują zasady geometrii euklidesowej. W celu opisania ogólnych geometrii trzeba było wyzwolić się od koncepcji owego osadzenia, co wiązało się ze sporym wyzwaniem w zakresie matematycznego myślenia abstrakcyjnego. Udało się to Gaussowi w 1827 roku w jego słynnym dziele o rozważaniach arytmetycznych *Disquisitiones Arithmeticae*. Potrafił on opisać geometrię ogólnych dwuwymiarowych płaszczyzn o dowolnej krzywiźnie w sposób wewnętrzny, korzystając wyłącznie ze współrzędnych zdefiniowanych na tej powierzchni, bez konieczności osadzania ich w trójwymiarowej przestrzeni. Dla płaszczyzn osadzonych wprowadził on wielkość pojęciowo niezależną od wewnętrznej krzywizny, nazywaną krzywizną zewnętrzną, która ma więcej wspólnego z tym, jak postrzegamy płaszczyznę z zewnątrz. W celu lepszego zrozumienia tego zagadnienia wyobraźmy sobie powierzchnię kartki papieru. Posiada ona znikającą wewnętrzną i zewnętrzną krzywiznę. Gdy powierzchnia kartki zostanie zwinięta, tworząc walec, jej wewnętrzna geometria nie ulegnie zmianie. Pozostanie ona nadal płaska, a suma kątów trójkąta wyniesie 180°. Zewnętrzna krzywizna jednak nie znika, lecz odpowiada za to, że zamiast płaskiej powierzchni widzimy walec. Gaussowi udało się udowodnić interesujący związek między wewnętrzną a zewnętrzną krzywizną, który nazwał *Theorema egregium*, czyli twierdzeniem wybornym. Głosi ono w skrócie, że z powodu trójwymiarowej przestrzeni euklidesowej określone składniki zewnętrznej i wewnętrznej krzywizny muszą się równoważyć. Powrócimy do tego twierdzenia później w kontekście grawitacji kwantowej (część 5).

Aby móc zapisać geometrycznie przestrzeń czterowymiarową, Einstein potrzebował uogólnienia wyników Gaussa do wyższych wymiarów. Pierwszy i najważniejszy krok w tym kierunku wykonał Georg Friedrich Bernhard Riemann (1826–1866). W swoim słynnym wykładzie habilitacyjnym, wygłoszonym 10 czerwca 1854 roku, mówił on w obecności wiekowego już Gaussa o hipotezach leżących u podstaw geometrii (*Über die Hypothesen, welche der Geometrie zu Grunde liegen* – wykład opublikowano w 1868 roku). Dzień ten uznaje się za decydujący dla uogólnienia teorii płaszczyzn Gaussa do dowolnych wymiarów, co nazwano później na jego cześć geometrią Riemanna. Również w tym przypadku najważniejszą kwestią jest możliwość matematycznego opisania przestrzeni bez konieczności odnoszenia się do przestrzeni osadzenia i związanych z tym poglądów.

Riemann poruszył także kwestię, która miała potem okazać się istotna w kontekście teorii względności, a mianowicie związek między geometrią a doświadczeniem. Riemann potrafił sobie bowiem wyobrazić, że istnieją takie zależności geometryczne, choćby w mikroskopijnym wymiarze, które nie odpowiadają już tylko płaskiej przestrzeni. Swój wykład zakończył słowami: „Kwestie te można rozstrzygnąć wyłącznie, opierając się na sprawdzonym doświadczalnie stanowisku Newtona, zastępując je stopniowo faktami, których nie można wyjaśnić za pomocą tego stanowiska (...). Prowadzi to w obszar innej nauki, fizyki, na grunt której natura tego spotkania nie pozwala wkroczyć". Naturalnie Riemann nie mógł jeszcze przypuszczać, że brakowało tu jednego kluczowego elementu, czyli uwzględnienia czasu. Dopiero zastosowanie geometrii Riemanna w czterowymiarowej czasoprzestrzeni pozwoliło na sformułowanie ogólnej teorii względności.

Czy geometryczny stan rzeczy można w ogóle potwierdzić doświadczalnie? Słynny matematyk Henri Poincaré (1854–1912) przyjął w tej kwestii szczególne stanowisko, nazywane także konwencjonalizmem. Gdy pojawia się sprzeczność między teorią a doświadczeniem, można zmienić geometrię lub właściwe prawa fizyki. Zawsze udaje się zachować

ważność geometrii euklidesowej, jeżeli tylko właściwe prawa zostaną odpowiednio dostosowane – wybór geometrii jest jedynie konwencją. W swoim uznanym wykładzie na temat geometrii i doświadczenia, *Geometrie und Erfahrung* z 1921 roku, Einstein wprowadza rozróżnienie między ściśle aksjomatyczną geometrią a geometrią praktyczną. Z Poincarém zgadza się jedynie na płaszczyźnie geometrii aksjomatycznej. W fizyce ważna jest geometria praktyczna, którą można zmierzyć prawdziwymi zegarami i skalami. Na przykład w przypadku ustalenia, że promienie świetlne poruszają się po prostych, sensowne wydaje się pytanie, czy świat obserwowany ma naturę euklidesową, czy też nie. Bez takiego rozumowania, zdaniem Einsteina, sformułowanie teorii względności nie byłoby możliwe. Częścią fizyki może być tylko tak rozumiana geometria. Ten punkt widzenia prezentowany jest także w niniejszej książce.

Czasoprzestrzeń musi być zatem opisana środkami geometrii Riemanna. Ponadto konieczne są również równania fizyczne, które określą dozwolone geometrie. Po żmudnej pracy Einstein zaprezentował słynne równania pola grawitacyjnego na posiedzeniu Królewsko-Pruskiej Akademii Nauk w Berlinie 25 listopada 1915 roku. Równania te w skomplikowany sposób łączą pole grawitacyjne z zawartością materii czasoprzestrzeni. W odpowiednim przypadku granicznym funkcję przybliżenia pełni ponownie zasada Newtona. Teoria Einsteina nie jest już jednak teorią oddziaływania na odległość, ponieważ oddziaływania rozprzestrzeniają się maksymalnie z prędkością światła. Nazwa „ogólna teoria względności" wiąże się z pomysłem, aby wszystkie układy odniesienia traktować jako równoważne. W odróżnieniu od szczególnej teorii nie ma już układów doskonałych, takich jak układy inercjalne. Pole grawitacyjne jest właściwością geometrii przestrzeni i czasu, a zatem geometria jest dynamiczna. Przestrzeń i czas nie są już absolutne, lecz wraz z materią uczestniczą w zjawiskach fizycznych we wszechświecie. W związku z tym stosowniejszą nazwą dla tej teorii byłaby *geometrodynamika* – analogicznie do elektrodynamiki.

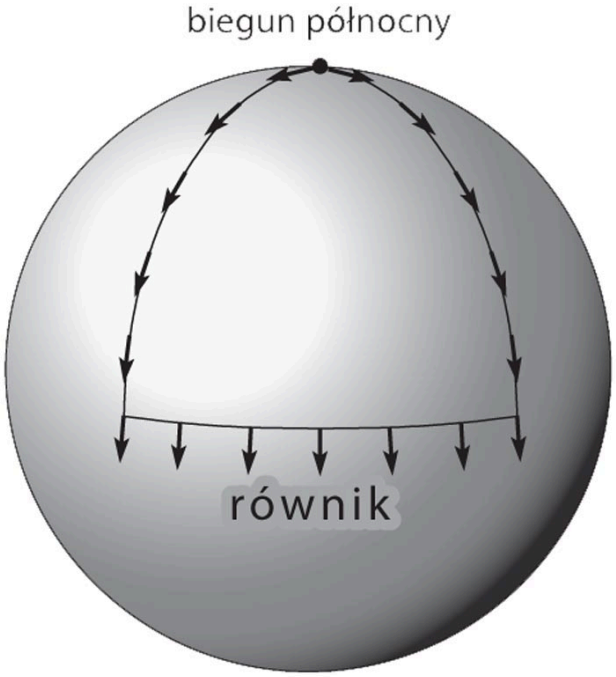
W ogólnej teorii względności Einsteina bezwładność i ciężkość są identyczne, gdyż opisują je te same wielkości matematyczne. Siły bezwładności są siłami odśrodkowymi, które odczuwamy na przykład na karuzeli. W teorii Newtona siły bezwładności różnią się wyraźnie od sił właściwych (jak choćby siły grawitacji). W teorii Einsteina takiego podziału już nie ma, a bezwładność jest w niej tylko określonym aspektem grawitacji.

Które z wielkości matematycznych są decydujące w geometrii Riemanna? Głównym pojęciem jest odległość między dwoma punktami. W geometrii euklidesowej przywołuje się w tym kontekście znane twierdzenie Pitagorasa. Jego uogólnienie prowadzi w geometrii Riemanna do pojęcia *metryki*, która podaje odległość między dwoma sąsiadującymi punktami, a na jej podstawie można z kolei ustalić odległość między dowolnymi punktami. Z metryki wynika także wielkość, która opisuje właściwą *krzywiznę* czasoprzestrzeni. Aby lepiej to zrozumieć, wyobraźmy sobie powierzchnię dwuwymiarową, na przykład powierzchnię kuli. Proste, czyli geodezyjne, są dane na tej powierzchni przez duże okręgi dzielące powierzchnię kuli na dwie części równej wielkości. Jeśli przyjrzymy się trójkątowi, jak na rysunku 1.9, złożonemu z odcinków prostych, możemy stwierdzić, że suma kątów jest większa niż 180°. Wielkość wykraczająca poza 180° jest właśnie miarą krzywizny. Istoty dwuwymiarowe, które mogą poruszać się tylko po powierzchni kuli, mogłyby dzięki temu stwierdzić, że ich geometria nie jest euklidesowa. Poza obliczeniem sumy kątów w trójkącie istoty te mogłyby dokonać także innych pomiarów geometrycznych.

Innym ważnym pojęciem w geometrii jest przesunięcie równoległe. Oznacza ono, że „wektor" przesuwany wzdłuż krzywej zawsze zachowuje ustalony kierunek, czyli nie podlega obrotowi. W przestrzeni euklidesowej to pojęcie jest oczywiste. Poznanie, jak w sposób ogólny można opisać przesunięcie równoległe w przestrzeni zakrzywionej, zawdzięczamy w znacznej mierze pracy włoskiego matematyka Tullia Leviego-Civity (1873–1941) z 1917 roku. Einstein był nią tak zafascynowany, że

2 sierpnia 1917 roku napisał do włoskiego uczonego: „To musi być wspaniałe uczucie galopować wśród pięknej przyrody na koniu prawdziwej matematyki, podczas gdy tacy jak my muszą iść pieszo”. Główną wielkością służącą do opisu przesunięcia równoległego jest tak zwana *spójność*, określana również mianem pola prowadzącego lub pola przesunięcia. Wskazuje ona, w jaki sposób, w wyniku przesunięcia równoległego, można dotrzeć od jednego punktu do drugiego, tym samym uzasadniając „spójność” między różnymi punktami. W geometrii Riemanna „spójność” można obliczyć na podstawie metryki, jednak może być ona także niezależna od niej. Na dwuwymiarowym przykładzie powierzchni kuli można zaobserwować przesunięcie równoległe wektora wzdłuż trójkąta. Na rysunku 1.9 ilustruje to niewielka strzałka przemieszczająca się wzdłuż sferycznego trójkąta.

Mimo że wektor zawsze zachowuje swój tymczasowy kierunek, po przebytym obiegu nie wskazuje on już w kierunku wyjściowym. Można wykazać, że zmiana kierunku jest miarą krzywizny płaszczyzny, którą obiega strzałka. W płaskiej przestrzeni krzywizna zanika, a kierunek nie zmienia się. Dla Einsteina spójność była wielkością nadrzędną, bardziej znaczącą niż metryka.



1.9. Wektor przesunięty równolegle względem wyznaczonego sferycznego trójkąta nie powraca do swojego kierunku wyjściowego

Pozostańmy jeszcze przez chwilę przy tym prostym przykładzie powierzchni kuli. Dwie istoty, które wyruszą z równika i pomaszerują w miarę możliwości prosto na północ wzdłuż wielkiego okręgu, spotkają się na biegunie północnym. Fakt ten można zinterpretować tak, że oddziałuje między nimi jakaś tajemnicza siła, która ich do siebie przyciąga. Takie byłoby stanowisko Newtona. W rzeczywistości jest to jednak tylko efekt geometrii, związany z faktem, że istoty te poruszają się po płaszczyźnie zakrzywionej. To z kolei stanowisko Einsteina. Analogiczną sytuacją w czasoprzestrzeni jest względne przyspieszenie lokalnych układów inercjalnych. W ogólnej teorii względności grawitacja *odpowiada* geometrii.

Wektor prowadzony po zakrzywionej przestrzeni wzdłuż zamkniętej krzywej doświadcza zmiany kierunku, lecz jego długość pozostaje taka sama. Sytuację taką opisuje ogólna teoria względności. Według matematyka Hermanna Weyla (1885–1955) takie zróżnicowane

traktowanie kierunku i długości było niezadowalające. W związku z tym w 1918 roku sformułował on uogólnioną teorię względności, która również przewidywała zmianę długości. Teoria ta miała w rzeczywistości pozwalać na dowolne zmiany skali („zmiany cechowania"), co dodatkowo umożliwiało opisanie elektromagnetyzmu. Niestety przyroda nie stosuje się do teorii Weyla. Jak zauważył Einstein, zgodnie z teorią Weyla bieg zegara musiałby być zależny od całej prehistorii zegara (jego linii świata). Nie istniałyby wtedy żadne pierwiastki chemiczne z liniami widmowymi o określonej częstotliwości, ponieważ każda częstotliwość jako proces okresowy jest zegarem. Mimo to praca Weyla nie poszła na marne. W artykule *Elektron und Gravitation* z 1929 roku zastosował on swoją koncepcję w zmienionej formie w odniesieniu do mechaniki kwantowej. Koncepcja ta przyczyniła się do powstania tak zwanych *teorii cechowania*, które z kolei leżą u podstaw współczesnych fizycznych teorii mikroświata. Więcej na ten temat w dalszej części książki.

Niegdyś uważano, że ogólna teoria względności znosi podział na system heliocentryczny i geocentryczny, ponieważ wszystkie układy odniesienia są równoważne. Nie jest tak. Wprawdzie zarówno w teorii Newtona, jak i Einsteina można zastosować opis do spoczywającej Ziemi lub spoczywającego Słońca, jednak siły Newtona i krzywizny Einsteina są wielkościami niezależnymi od układów odniesienia. Słońce zdominowało zależności geometryczne w Układzie Słonecznym, do których planety muszą się dostosować (zob. następny rozdział). Zatem także w ramach ogólnej teorii względności można bez obaw powiedzieć, że Ziemia obraca się wokół własnej osi i wokół Słońca. Wspomniana już aberracja, jak i paralaksy gwiazd, które wynikają ze zmiany położenia Ziemi podczas obrotu wokół Słońca, mówią zgodnie: a jednak się kręci!

Teoria względności w Układzie Słonecznym

Zasada równoważności Einsteina stanowi trzon ogólnej teorii względności. Zakłada ona, że jeden (niewielkich rozmiarów) swobodnie spadający układ jest równoważny względem układu w przestrzeni, w której nie

oddziałuje siła ciężkości (porównaj z rys. 1.6). Analogicznie zatem nie można odróżnić lokalnie układu w polu grawitacyjnym od układu podlegającego przyspieszeniu w przestrzeni, w której nie oddziałuje siła ciężkości. Z tego wynika zrównanie grawitacji z geometrią przestrzeni czterowymiarowej. Równania pola Einsteina są więc potrzebne do określenia możliwych geometrii w zależności od materii.

Już sama zasada równoważności prowadzi do interesujących wniosków bez konieczności korzystania z równań pola. Rozważmy przykład przyspieszającej rakiety. W jej wnętrzu przesyłana jest fala świetlna ze źródła znajdującego się na ziemi w kierunku sufitu, gdzie umieszczono detektor. Z uwagi na to, że detektor z powodu przyspieszenia oddala się od napływających fal świetlnych, przedział czasu między nadejściem następujących po sobie grzbietów i dolin fal jest większy niż podczas opuszczania źródła. Zjawisko to nazwano *przesunięciem ku czerwieni*, ponieważ długość fali świetlnej zwiększa się, a światło czerwone wykazuje większą długość fali niż światło niebieskie. Ten efekt zmiany długości fali określa się mianem efektu Dopplera. W wersji akustycznej jest nam dobrze znany z codzienności. Gdy światło przemieszcza się w odwrotnym kierunku, z góry na dół, pędzi w stronę źródła i doznaje przesunięcia ku fioletowi, czyli zachodzi zjawisko skrócenia długości fali.

Z uwagi na to, że zasada równoważności postuluje lokalną równoważność przyspieszenia i pola grawitacyjnego, przesunięcie ku czerwieni musi nastąpić także wtedy, gdy wyślemy falę świetlną z powierzchni Ziemi na określoną wysokość. Efekt ten został już potwierdzony w latach sześćdziesiątych XX wieku w eksperymencie przeprowadzonym w niespełna dwudziestometrowej wieży na Uniwersytecie Harvarda mimo niewielkich wartości wynoszących około 10^{-15} dla względnej zmiany długości fali z niepewnością rzędu jednego procenta. Już na podstawie tego zaobserwowanego efektu przesunięcia ku czerwieni można powiedzieć, że czasoprzestrzeń nie może być płaska.

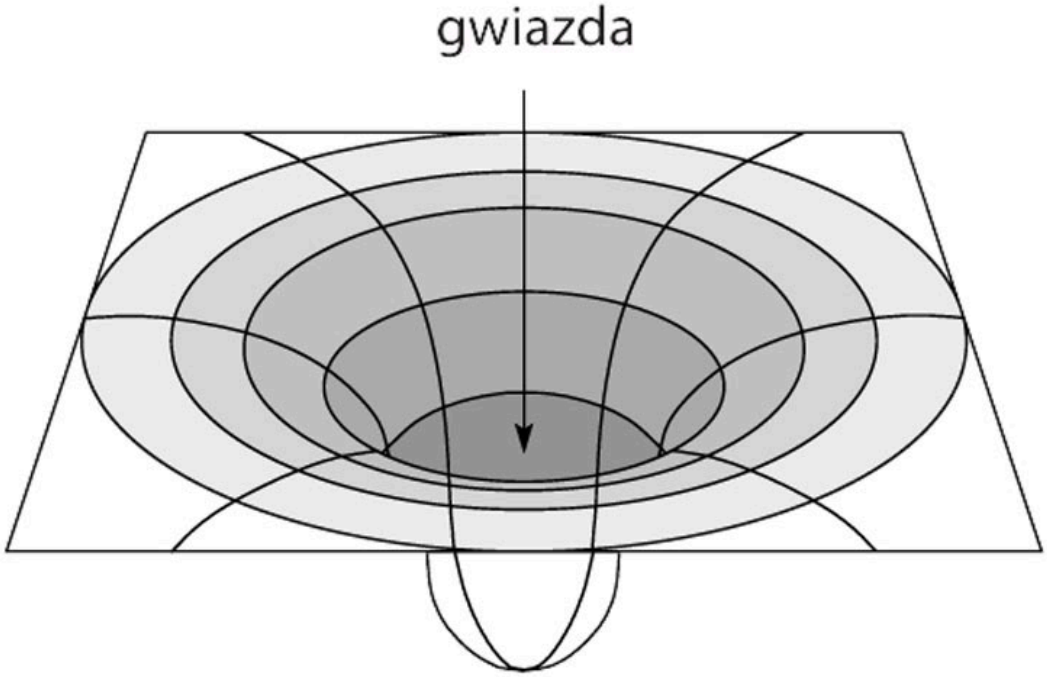
Fala świetlna jest tylko szczególnym przypadkiem ogólnego ruchu periodycznego. Z uwagi na to, że każdy proces periodyczny można

potraktować jak zegar, z zasady równoważności wynika, że ruch zegara zależy od miejsca w polu grawitacyjnym. Nie chodzi tu o wewnętrzny ruch zegara, który jest dla niego zawsze ten sam, niezależnie od tego, gdzie się znajduje (zakładając istnienie zegara idealnego), ale raczej o względny ruch dwóch zegarów o tej samej budowie, które znajdują się na różnej wysokości w polu grawitacyjnym. Zegar znajdujący się niżej chodzi względnie wolniej. Takie sformułowanie jest nieco niedbałe, lecz ogólnie stosowane. Podobnie jak w przypadku problemu bliźniaków, chodzi po prostu o to, że linie świata zegarów w czasoprzestrzeni mają różne długości. Rzeczywista różnica w ruchu zegarów powyżej Ziemi jest jednak niewielka. Z powodu grawitacyjnej dylatacji czasu zegar w Kolonii spóźnia się w porównaniu z zegarem o takiej samej budowie umieszczonym na szczycie Zugspitze o jedną milisekundę na sto lat.

Pomimo swojej znikomości efekt ten znalazł zastosowanie w naszym życiu codziennym. Obecnie istnieją systemy pozycjonowania, które pozwalają określić dowolne miejsce na Ziemi z dokładnością do jednego metra, a nawet dokładniej. Najstarszym z nich jest obsługiwany ze Stanów Zjednoczonych *Global Positioning System* (GPS). Istnieje również jego rosyjski odpowiednik GLONASS oraz europejski GALILEO, który ma zostać uruchomiony w 2013 roku. Sposób funkcjonowania tych systemów opiera się na szeregu okrążających Ziemię satelitów wyposażonych w zegary atomowe. Satelity wysyłają sygnały elektromagnetyczne zawierające informacje o miejscu i czasie nadania sygnału. Gdy obserwator odbierze sygnały od co najmniej czterech satelitów, może ustalić własną pozycję (trzy współrzędne) i własny czas. Z uwagi na niezwykle wysoką dokładność konieczne jest uwzględnienie zjawisk szczególnej i ogólnej teorii względności. Bieg zegara podlega wpływowi zarówno prędkości zegara, jak i miejsca w polu grawitacyjnym. Prędkość sygnałów jest zawsze równa prędkości światła c, niezależnie od prędkości satelitów i odbiorcy. Gdyby nie uwzględniano tych zjawisk, już w ciągu jednego dnia pojawiłyby się błędne wskazania rzędu wielu kilometrów, a urządzenie stałoby się całkowicie bezużyteczne. Każdy użytkownik GPS potwierdza tym samym pośrednio, że czasoprzestrzeń jest zakrzywiona.

Grawitacyjny wpływ na bieg zegara wynika zatem bezpośrednio z zasady równoważności. Inne efekty wymagają poszukiwania rozwiązań równań pola Einsteina. Jedno z najważniejszych, dokładnych rozwiązań opracowano niedługo po ukończeniu ogólnej teorii względności. Powołany do służby wojskowej niemiecki astronom Karl Schwarzschild (1873–1916) na froncie wschodnim w 1916 roku odkrył czasoprzestrzeń, która opisuje pustą przestrzeń zewnętrzną sferycznie symetrycznego rozkładu mas. Na cześć odkrywcy mówi się o rozwiązaniu Schwarzschilda. Jedynym istotnym parametrem jest masa ciała. Rozwiązanie to jest szczególnie istotne dla opisu Układu Słonecznego, którego geometria jest w znacznym stopniu uzależniona od dominującej masy Słońca.

Czterowymiarowa czasoprzestrzeń jest trudna do zilustrowania. W przypadku rozwiązania Schwarzschilda należy zatem skonfrontować się z dwoma stanami rzeczy. Opisana przez nie geometria jest statyczna, czyli niezależna od czasu, i jest sferycznie symetryczna. Wystarczy więc przyjrzeć się stałemu punktowi w czasie oraz stałemu przekrojowi pozostałej trójwymiarowej przestrzeni – redukuje to czterowymiarową czasoprzestrzeń do dwóch wymiarów przestrzeni. Aby to zobrazować, płaszczyznę tę można umieścić w trójwymiarowej płaskiej przestrzeni pomocniczej. Wynik tego zabiegu jest przedstawiony na rysunku 1.10, przy czym należy pamiętać, że rozwiązanie Schwarzschilda obowiązuje jedynie na zewnątrz zaznaczonej tam gwiazdy (Słońca). W związku z tym na brzegu Słońca dołączono rozwiązanie, które opisuje jego wnętrze. W gruncie rzeczy rysunek przypomina trampolinę z wgłębieniem pośrodku, utworzonym przez masę Słońca. Oczywiście porównanie to nie jest precyzyjne, ponieważ obserwacja czasoprzestrzeni Schwarzschilda z zewnątrz nie jest możliwa, mimo to można dzięki temu zobrazować kilka efektów geometrycznych.

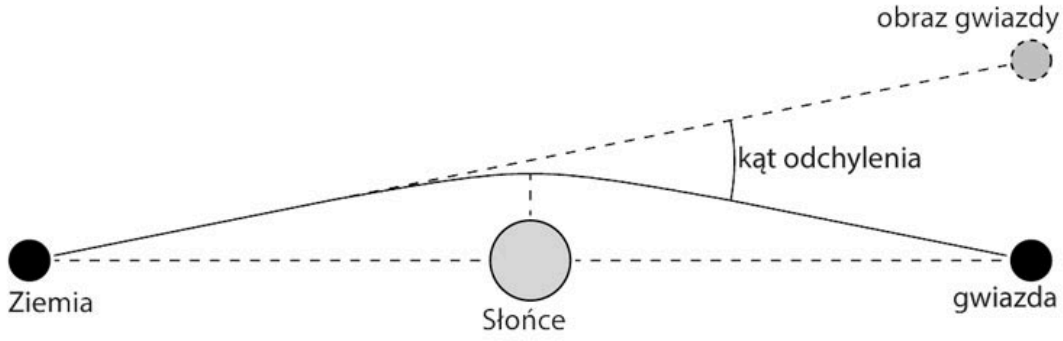


1.10. Geometria w otoczeniu kulistej gwiazdy

Początkowo na brzegu Słońca pojawia się odchylenie światła – zjawisko ważne dla historii teorii względności. Chociaż pojęcie „odchylenie” weszło na stałe do języka, to ściśle rzecz ujmując, nie jest ono poprawne. Oczywiście także w czasoprzestrzeni Schwarzschilda światło biegnie po geodezyjnych, czyli możliwie najprostszych liniach. Odchylenie odnosi się zatem do porównania do płaskiej przestrzeni. Promień świetlny jest obserwowany raz przy obecnym, a raz przy nieobecnym Słońcu. Różnicą jest odchylenie światła (porównaj z rys. 1.11).

Zjawisko to udało się zmierzyć już przy pierwszej próbie. W tym celu porównano płyty fotograficzne przedstawiające ten sam region gwiazd przy obecności i nieobecności Słońca. Oczywiście pomiar w pobliżu Słońca był możliwy jedynie przy jego całkowitym zaćmieniu. Aby dokonać stosownych obserwacji, podczas zaćmienia 29 maja 1919 roku wysłano dwie ekspedycje pod kierownictwem sir Arthura Stanleya Eddingtona (1882–1944) – jedną do Sobral w Brazylii, a drugą na Wyspę Książęcą, na zachodnim wybrzeżu Afryki. Wyniki potwierdziły, choć

z pewną niepewnością, prognozy teorii względności, a jej twórcę uczyniły sławnym na całym świecie.



1.11. Odchylenie światła wskutek oddziaływania Słońca

Obecnie za pomocą sieci radioteleskopów mierzy się odchylenie sygnałów radiowych, które pochodzą z pozagalaktycznych źródeł emisji fal radiowych (np. kwazarów). Dzięki temu udało się potwierdzić prognozowane odchylenie z niepewnością wynoszącą mniej niż jedną dziesiątą promila.

Droga światła (lub ogólnie promieniowania elektromagnetycznego) jest w okolicach Słońca dłuższa, ponieważ musi podążać po zakrzywionej przestrzeni. W związku z tym światło również biegnie dłużej (w porównaniu z płaską przestrzenią). To „opóźnienie biegu" pozwala na jeszcze jeden, nieco dokładniejszy test teorii względności. Zjawisko to można było sprawdzić w 2002 roku dzięki wymianie sygnałów z sondą Cassini, która znajdowała się w drodze do Saturna. Niepewność wyniosła mniej niż cztery setne promila.

Zjawisko odchylenia światła jest istotne także poza Układem Słonecznym. Światło odległych galaktyk lub kwazarów może na swojej drodze ku Ziemi, mijając galaktykę bądź gromadę galaktyk, ulec odchyleniu. Takie obiekty powodujące odchylenie określa się mianem soczewek grawitacyjnych. Dzięki temu na niebie można zaobserwować zwielokrotnione obrazy odległych obiektów. Z uwagi na to, że światło

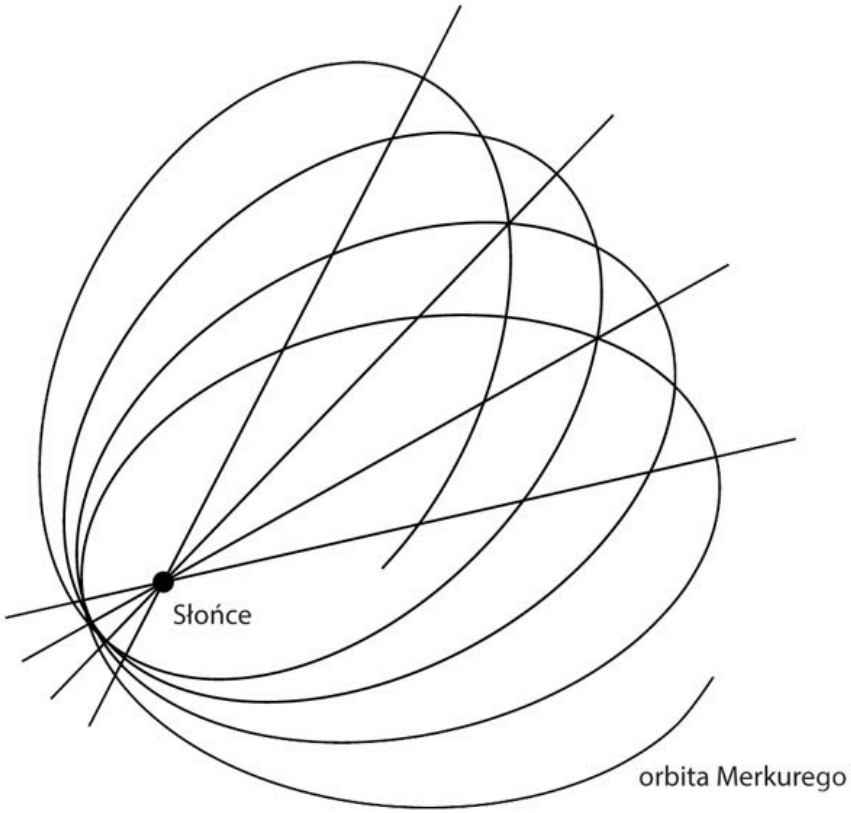
obiektów podlegających efektowi soczewkowania grawitacyjnego pokonuje duże odległości i zostało wysłane przed wieloma milionami lat, można dzięki nim niejako zmierzyć wszechświat. Teorię soczewek grawitacyjnych Einstein opisał w swoim notatniku już w 1912 roku, lecz opublikował ją dopiero w 1936 roku, gdyż z powodu niewielkiego rozmiaru tego efektu nie wierzył w możliwość jego obserwacji. Trzeba było poczekać do 1979 roku, zanim rzeczywiście stało się to możliwe.

Dzisiaj soczewki grawitacyjne należą do astronomicznej codzienności. Jeżeli źródło leży bezpośrednio za soczewką grawitacyjną, to odchylone światło, z uwagi na zjawisko symetrii, tworzy pierścień wokół soczewki, czyli tak zwany pierścień Einsteina. Rysunek 1.12 przedstawia rzadki przypadek podwójnego pierścienia Einsteina, który został odkryty w 2007 roku. Wówczas światło *dwóch* odległych galaktyk zostaje rzeczywiście odchylone przez soczewkę grawitacyjną. Dzięki efektowi soczewki grawitacyjnej możliwe stało się również wyśledzenie planet spoza naszego Układu Słonecznego.

Jednym z największych osiągnięć mechaniki Newtona było wyjaśnienie praw Keplera dotyczących ruchu planet. Planety poruszają się po orbitach eliptycznych, w których ognisku znajduje się Słońce. Twierdzenie to jest poprawne, dopóki nie jest brany pod uwagę wpływ innych planet. W momencie uwzględnienia tych wpływów dochodzi do zmiany orbity, którą można scharakteryzować na przykład przez precesję peryhelium, czyli najbliższego Słońcu punktu orbity. Już w XIX wieku sztuka mechaniki nieba była tak dobrze rozwinięta, że możliwe było poprawne obliczenie wpływu innych planet i tym samym określenie precesji peryhelium. W przypadku Merkurego, najbliższej Słońcu planety, rozbieżność wobec obserwacji wyniosła około 43 sekund kątowych na 100 lat, co wielu osobom spędzało sen z powiek. Taka rozbieżność stanowi około ośmiu procent całej precesji peryhelium.

1.12. Zdjęcie podwójnego pierścienia Einsteina wykonane za pomocą teleskopu Hubble'a

Według ogólnej teorii względności planeta, także bez wpływu innych planet, nie porusza się już po elipsie. Dopóki jednak nie zbliży się ona zbytnio do Słońca, odchylenie od elipsy jest niewielkie, ponieważ krzywizna powodowana przez Słońce zmniejsza się wraz z odległością od niego. Małe odchylenie wyraża się między innymi w precesji peryhelium, która prowadzi do orbity w kształcie rozety zamiast elipsy (rys. 1.13). W przypadku najbliższego Słońcu Merkurego efekt ten jest najbardziej wyraźny.



1.13. Precesja peryhelium planet

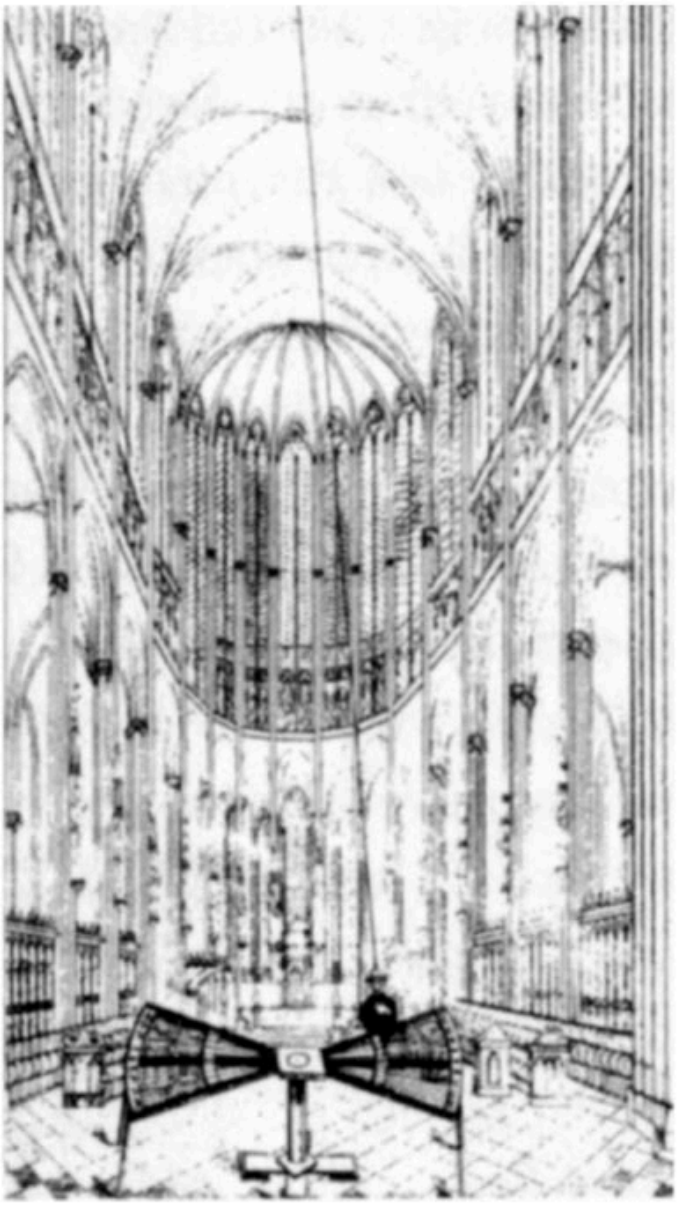
Einstein był niezmiernie szczęśliwy, gdy odkrył, że jego teoria może wyjaśnić te brakujące 43 sekundy kątowe przypadające na każde stulecie. Podobno nazwał tę chwilę najszczęśliwszą w swoim życiu. 21 grudnia 1915 roku pisał do swego długoletniego przyjaciela Michele Besso, aby ten przeczytał rozprawy, bo jest w nich ostateczne wyjście z tej, jak to określił, biedy, a najbardziej pomyślną informacją jest zgodność precesji peryhelium. Obrót o 43 sekundy kątowe na 100 lat odpowiada pełnemu ruchowi po okręgu co 30 000 stuleci. Pomimo niewielkiej skali efekt ten jest jednak istotny dla naszego zrozumienia czasu i przestrzeni. Tymczasem został on już potwierdzony z niepewnością rzędu około trzech promili.

To samo dotyczy kolejnego efektu. Także i on jest niezmiernie znikomy jak na warunki panujące w Układzie Słonecznym, ale jeszcze lepiej

uwidacznia dynamiczną naturę czasoprzestrzeni, niż czyni to precesja peryhelium. Według teorii Newtona tylko masy wytwarzają pole grawitacyjne. Według teorii Einsteina strumienie mas mogą funkcjonować dodatkowo jako źródło. Jest to sytuacja analogiczna do elektromagnetyzmu, w przypadku którego same ładunki wytwarzają pole elektryczne, natomiast strumienie ładunków – pole magnetyczne. Ze względu na podobieństwo do magnetyzmu efekt ten nazywany jest efektem grawitomagnetycznym. Ważnym przypadkiem szczególnym jest rotacja ciała kulistego, na przykład Ziemi. Wskutek występowania rotacji dochodzi do przyciągania lokalnych układów inercjalnych znajdujących się w jej pobliżu. W rzeczywistości to pusta czasoprzestrzeń ulega oddziaływaniu i sama oddziałuje na lokalne układy inercjalne.

Czym charakteryzuje się taki efekt? Na początek przyjrzyjmy się klasycznemu instrumentowi, jakim jest wahadło Foucaulta, które dostarcza dowodu na to, że Ziemia się kręci. Chodzi o proste wahadło o odpowiedniej długości, którego płaszczyzna drgań jest stale osadzona w przestrzeni, czyli nie zmienia się w stosunku do odległych gwiazd stałych (*International Celestial Reference System*). Z uwagi na to, że Ziemia porusza się względem tego systemu odniesienia, obserwator na Ziemi dostrzega ruch wahadła względem Ziemi. Historyczny eksperyment odbył się w 1851 roku w Panteonie w Paryżu i tam rozpoczął swój triumfalny pochód przez świat. Eksperyment taki przeprowadzono również w katedrze św. Piotra i Najświętszej Marii Panny w Kolonii w 1852 roku (zob. rys. 1.14). Galileusz byłby z pewnością zachwycony perspektywą przeprowadzenia w kościele eksperymentu mającego dowieść, że Ziemia się kręci.

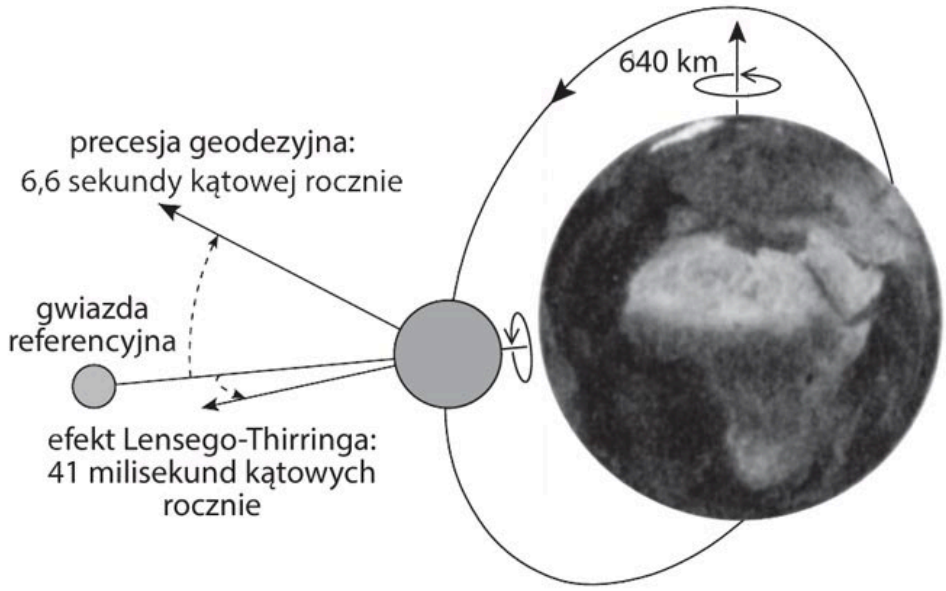
Według teorii Einsteina obracająca się Ziemia przyciąga lekko przestrzeń ze swojego otoczenia, a tym samym także płaszczyznę drgań wahadła. Na cześć dwóch austriackich fizyków, Hansa Thirringa (1888–1976) i Josepha Lensego (1890–1985), którym udało się obliczyć ten efekt przy wykorzystaniu ogólnej teorii względności, nazwano go efektem Lensego-Thirringa.



1.14. Szkic przedstawiający eksperyment z użyciem wahadła Foucaulta w katedrze w Kolonii w 1852 roku

Wahadło nie jest optymalnym wyborem w przypadku bezpośredniego testu, ponieważ nie znajduje się w układzie inercjalnym (układzie swobodnie spadającym), lecz spoczywa na ziemi. W związku z tym do pomiaru używa się precesji żyroskopu, który znajduje się w sondzie kosmicznej, a tym samym swobodnie spada. (Pod pojęciem precesji rozumie się zmianę kierunku osi obracającego się ciała). Żyroskop taki

ulega wtedy precesji z określoną prędkością kątową względem *International Celestial Reference System*. W przeciwieństwie do teorii Newtona precesja taka nie wymaga żadnych sił zewnętrznych. W rzeczywistości owe siły byłyby potrzebne do zatrzymania żyroskopu względem odległych gwiazd stałych. Po trwających dziesiątki lat przygotowaniach w kwietniu 2004 roku doszło do startu sondy kosmicznej o nazwie „Gravity-Probe B", na pokładzie której umieszczono takie żyroskopy. Sonda poruszała się po okołobiegunowej orbicie zamkniętej na wysokości 640 kilometrów nad Ziemią. Dla tej wysokości precesja żyroskopu względem odległych gwiazd stałych wynosi na podstawie efektu Lensego-Thirringa tylko około 41 milisekund kątowych na rok, co odpowiada pełnemu obiegowi w ciągu około 31 milionów lat! Na rysunku 1.15 przedstawiono zależności geometryczne. Zaznaczono także wartościowo większy efekt precesji geodezyjnej, następujący wokół innej osi. Zachodziłaby ona również przy nieruchomej Ziemi, a wynika to z tego, że w zakrzywionej przestrzeni oś żyroskopu zmienia po drodze swój kierunek. Analiza danych dotyczących tego eksperymentu nadal trwa, a jej skuteczne ukończenie potwierdziłoby jeden z najbardziej fascynujących efektów ogólnej teorii względności.



1.15. Efekt Lensego-Thirringa i precesja geodezyjna

Efekt Lensego-Thirringa jest często łączony z zasadą Macha. Wywodzi się ona od austriackiego fizyka i filozofa Ernsta Macha (1838–1916) i postuluje eliminację wszystkich elementów absolutnych w teorii fizycznej. Mach miał na myśli głównie przestrzeń absolutną i czas absolutny Newtona. W rzeczywistości jednak zasada ta jest sformułowana w sposób tak mglisty, że dopuszcza różnorodne interpretacje. Sam Einstein określił zasadę Macha jako ważny przyczynek do teorii względności. Pierwotnie rozumiał ją tak, że geometria czasoprzestrzeni (metryka) powinna być jednoznacznie określana przez rozkład materii. *Nie* odpowiada to jednak ogólnej teorii względności, co Einstein stwierdził w późniejszym czasie. Metryka jest obecna we wszystkich równaniach dotyczących materii i dlatego nie może być określana przez materię. Interpretacja zasady Macha przez Einsteina nie może zatem być poprawna, ponieważ zbyt mocno opiera się na zasadzie oddziaływania na odległość Newtona. Według teorii względności pole grawitacyjne posiada własne lokalne stopnie swobody, które nie są ustalane przez źródła. Jest to główna myśl teorii pola. To lokalne pola grawitacyjne są odpowiedzialne za

lokalną bezwładność (a nie, jak twierdził Mach, odległe gwiazdy stałe). Dokładniej rzecz ujmując, miarodajną wielkością jest opisana powyżej „spójność", czyli matematyczna wielkość odpowiedzialna za przesunięcie równoległe w zakrzywionej przestrzeni. Spójność ta istnieje nawet wtedy, gdy czasoprzestrzeń jest płaska. Właśnie to miał na myśli Einstein, mówiąc o równoważności istoty bezwładności i grawitacji. Podczas gdy u Newtona oba te fenomeny są mocno rozgraniczone (siły bezwładności są definiowane przez przestrzeń absolutną, a grawitacyjne przez przyciąganie masy), u Einsteina tworzą one dwa aspekty tej samej rzeczy, czyli geometrii czasoprzestrzeni. W związku z tym siła bezwładności, którą odczuwamy podczas przejażdżki na karuzeli, jest aspektem grawitacji.

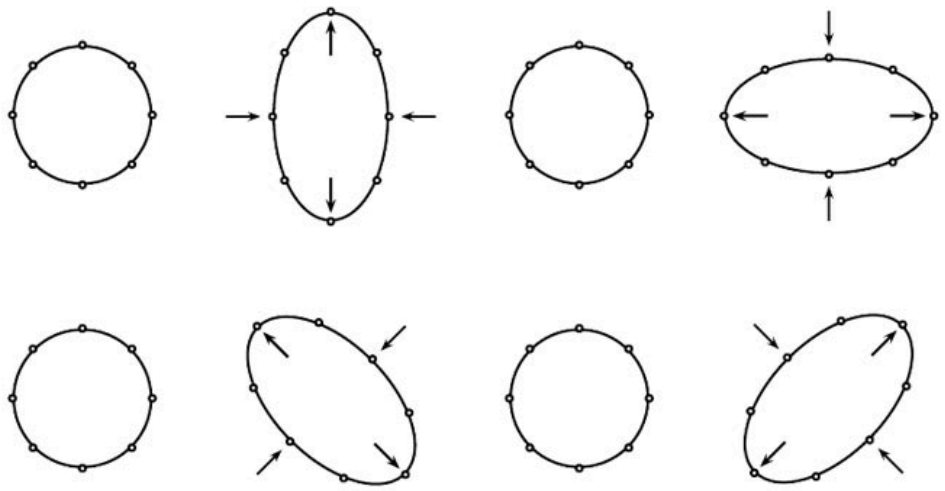
Fale grawitacyjne

John Archibald Wheeler (1911–2008) opisał istotę równań pola Einsteina w taki sposób, że materia mówi czasoprzestrzeni, jak ma się zakrzywiać, a czasoprzestrzeń materii, jak ma się poruszać. Do tego można jeszcze dodać, że czasoprzestrzeń mówi sama sobie, jak ma się zakrzywiać. Pole grawitacyjne posiada bowiem własne dynamiczne stopnie swobody, które mogą się rozprzestrzeniać niezależnie od materii. Już w 1916 roku Einstein przewidywał istnienie fal grawitacyjnych. Jego równania dopuszczają istnienie rozwiązań falowych – analogicznie do odkrytych przez Hertza fal elektromagnetycznych. Podobnie jak tamte poruszają się one z prędkością światła i drgają prostopadle do kierunku ich rozchodzenia się. Chodzi tu jednak nie o fale, które podobnie jak światło poruszają się w zadanej przestrzeni, lecz o okresowe zakłócenia pustej przestrzeni, przypominające bardziej fale wodne, gdzie woda symbolizuje przestrzeń.

Jak rozpoznać takie fale grawitacyjne? Gdy natrafiają one na swobodnie spadające masy testowe, powodują okresowe rozszerzanie i kurczenie odległości między masami, zmieniając tym samym geometrię przestrzeni oraz długości. Do tego niezbędne są co najmniej *dwie* masy, ponieważ jedna masa tylko spada swobodnie i w związku z tym nie daje możliwości

zaobserwowania zmiany odległości. Na rysunku 1.16 przedstawiono, jak pierścień takich mas testowych zmienia okresowo swój kształt wskutek prostopadłych uderzeń fal grawitacyjnych.

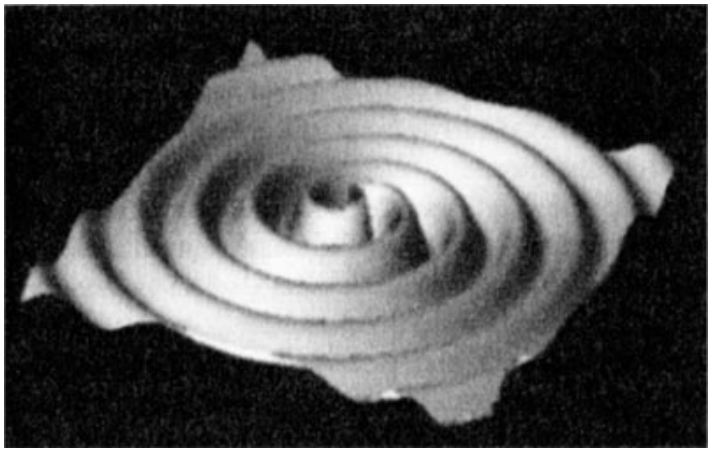
Fale grawitacyjne powstają, gdy rozkład mas zmienia się nagle w sposób asymetryczny, to znaczy nie kulisty. Może chodzić o kolaps wypalonej gwiazdy, który prowadzi do wybuchu supernowej, albo o stapianie się zwartych obiektów (gwiazd neutronowych, czarnych dziur), o drgania i rotacje gwiazd neutronowych lub o skomplikowane procesy we wczesnym wszechświecie, które zostawiają wieczny ślad w formie „pierwotnych fal grawitacyjnych". Na rysunku 1.17 pokazano przykład fali grawitacyjnej, emitowanej przez dwie krążące wokół siebie gwiazdy neutronowe.



1.16. Oddziaływanie fali grawitacyjnej, uderzającej pionowo w płaszczyznę papieru, na kółko cząsteczek kurzu. Dochodzi do okresowego kurczenia się i rozszerzania, przy czym oś w drugim rzędzie obrócona jest o 45°. Ogólna sytuacja powstaje przez nałożenie się na siebie obu przypadków.

Większość częstotliwości znajduje się w „obszarze akustycznym" kilkuset herców, czyli jest o wiele bardziej nudna niż fale świetlne czy radiowe. W związku z tym sygnały te można przedstawić bezpośrednio

jako sygnał słyszalny, którego dysonanse zrekompensują brak bezpośredniej słyszalności.



1.17. Dwuwymiarowy obraz fal grawitacyjnych, emitowanych przez dwie krążące wokół siebie gwiazdy neutronowe

Fakt, że fale grawitacyjne mogą transportować energię i wykonywać pracę, wynika już z prostego argumentu, który doskonale uwidacznia różnicę między teorią Einsteina a Newtona. U Newtona siła grawitacji działa natychmiastowo na dowolne odległości, natomiast u Einsteina oddziaływanie to może się rozprzestrzeniać najwyżej z prędkością światła. Przypatrzmy się dwóm ciałom A i B, które zbliżają się do siebie wskutek przyciągania grawitacyjnego. Na potrzeby naszego eksperymentu myślowego załóżmy, że masy te w wyniku zderzenia z mechanizmem stalowych sprężyn zostają od siebie odepchnięte i zawracają, a następnie ponownie się do siebie zbliżają pod wpływem działania grawitacji. Podczas ruchu na zewnątrz A odczuwa przyciąganie B, zupełnie jakby oba znajdowały się bliżej siebie – oddziaływanie grawitacji potrzebuje skończonego czasu, aby dotrzeć od B do A. Przyciągająca siła B jest zatem większa niż podczas natychmiastowego oddziaływania grawitacji. W odwrotnej sytuacji ciało A w drodze do wewnątrz odczuwa siłę B, jakby oba były od siebie bardziej oddalone – siła jest więc mniejsza. W związku z tym w drodze na zewnątrz wykonywana jest większa praca, niż w drodze do wewnątrz jest uzyskiwana. Z uwagi na to, że energia

ostatecznie zostaje zachowana, część energii musi opuścić lokalny obszar mas – chodzi tu właśnie o energię fal grawitacyjnych, które uciekają stamtąd z prędkością światła.

Jak wspomniano powyżej, istnienie fal grawitacyjnych można potwierdzić dzięki ich oddziaływaniu na masy testowe. Odległości między takimi masami okresowo rozciągają się i kurczą. Wielkość tej deformacji zmniejsza się odwrotnie proporcjonalnie do odległości od powodującego ją źródła. To o wiele mniejszy ubytek niż w przypadku statycznego prawa powszechnego ciążenia Newtona, według którego siła zmniejsza się odwrotnie proporcjonalnie do *kwadratu* odległości. Mimo to oczekiwana deformacja jest niewielka, gdyż mocne źródła znajdują się w odległościach astronomicznych, na przykład odcinek o długości jednego kilometra skraca się lub wydłuża o wartość odpowiadającą wielkości jądra atomu bądź mniejszą, to jest 10^{-13} centymetra, czyli jedną stutysięczną średnicy atomu. Czy istnieje w ogóle możliwość pomiaru tak mikroskopijnych zmian odległości? Pionierem takich pomiarów był amerykański fizyk Joseph Weber (1919–2000), który chciał potwierdzić istnienie fal grawitacyjnych, dowodząc, że wprawiają one w drania aluminiowy walec. Niestety czułość urządzenia była zbyt mała, aby uzyskać odpowiedni dowód. Obecnie na całym świecie istnieją cztery takie urządzenia, które dostarczają wyniki ze 100 tysięcy razy większą czułością. Na dowód istnienia fal wciąż jednak czekamy.

Największe nadzieje wiąże się z tak zwanymi interferometrami laserowymi. Na całym świecie jest obecnie pięć takich urządzeń odpowiedniej wielkości. Wykorzystuje się w nich interferencję światła. Promień lasera pada na półprzepuszczalną powierzchnię lustra i zostaje tam rozdzielony na dwie składowe promienia. Składowe promienia biegną w kierunku dwóch luster, tam zostają odbite i nałożone w detektorze, to znaczy doprowadzone do interferencji. Powstaje wzorzec interferencyjny jasnych i ciemnych miejsc, odpowiednio do wzajemnego wzmocnienia (grzbiet fali na grzbiet fali lub dolina fali na dolinę fali) bądź wygaszenia (grzbiet fali na dolinę fali) składowych fali. Wzorzec taki zależy od

względnej odległości dróg świetlnych. Gdy pojawia się jedna fala grawitacyjna, w typowy dla siebie sposób zmienia tę względną odległość, a tym samym również sam wzorzec interferencyjny. Oczywiście dowód taki może zostać uznany tylko wtedy, gdy wiele detektorów na Ziemi zmierzy ten sam sygnał. W przeciwnym razie nigdy nie byłoby pewności co do tego, czy jednak nie doszło do zakłócenia. Oprócz dwóch urządzeń zlokalizowanych w Stanach Zjednoczonych oraz jednego w Japonii i we Włoszech istnieje jeszcze jeden interferometr w Niemczech niedaleko Hanoweru. Urządzenie to zostało zbudowane w ramach niemiecko-brytyjskiego projektu GEO600 i rozpoczęło pracę w 2006 roku. Długość jego ramienia wynosi 600 metrów. W każdej chwili może dojść do namierzenia pierwszego jednoznacznego sygnału fal grawitacyjnych. Aby uniknąć zakłóceń, przede wszystkim pochodzenia sejsmicznego, planowane jest nawet umieszczenie interferometru laserowego w kosmosie. Projekt LISA (*Laser Interferometer Space Antenna*), zorganizowany wspólnie przez NASA i ESA, ma się składać z trzech satelitów, które utworzą trójkąt o długości boku wynoszącej pięć milionów kilometrów. Ich zadaniem będzie podążanie za Ziemią na jej drodze wokół Słońca. Promienie lasera będą wymieniane między satelitami. LISA ma zacząć działać w 2018 roku (lub nieco później). W połączeniu z ziemskimi urządzeniami dałoby to początek astronomii fal grawitacyjnych. Z uwagi na to, że fale grawitacyjne wskutek oddziaływania słabego pędzą bez przeszkód przez wszechświat, zawierają one informacje pochodzące z odległych zakątków galaktyki i dalekiej przeszłości, których nie potrafimy uzyskać w żaden inny sposób.

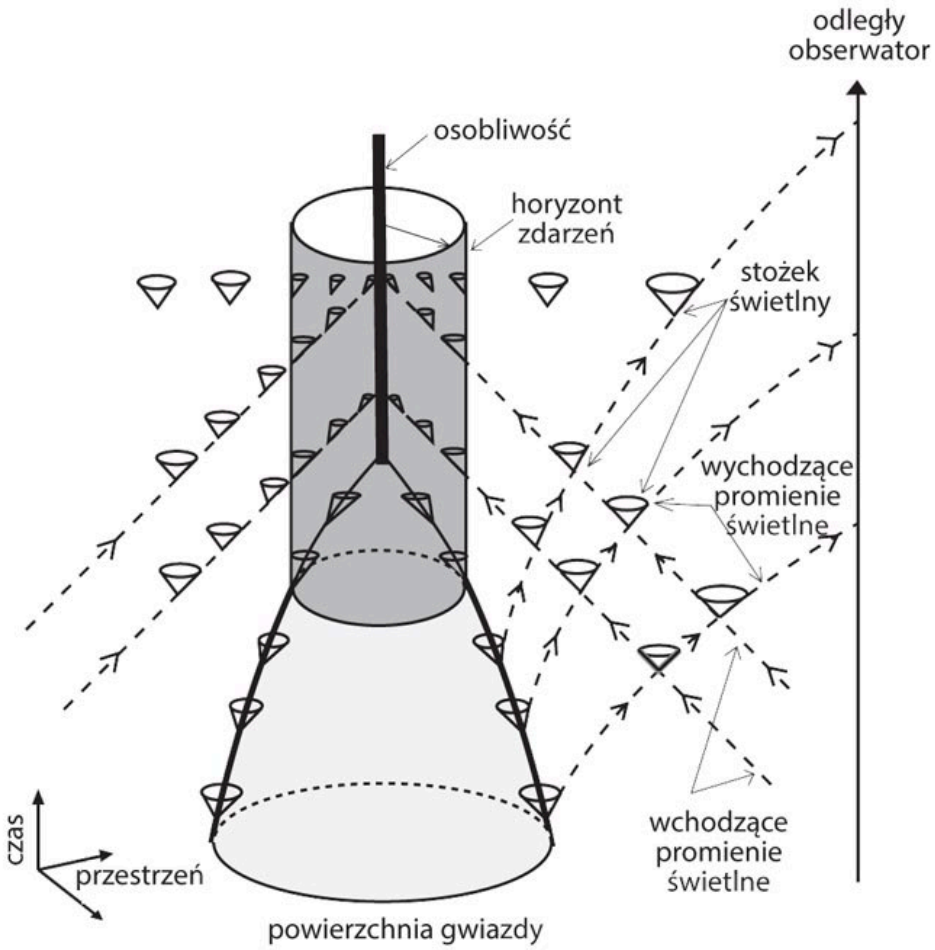
W rzeczywistości *udało się* już potwierdzić istnienie fal grawitacyjnych, choć tylko pośrednio. Według teorii Einsteina fale takie emituje wzajemnie okrążająca się para gwiazd podwójnych. Wskutek tego wynoszona jest energia, którą utracił układ gwiezdny. To z kolei prowadzi do skrócenia okresu orbitalnego – gwiazdy zbliżają się do siebie coraz bardziej i zaczynają coraz szybciej wirować wokół siebie. W sprzyjających warunkach można nawet zaobserwować to zjawisko. Istnieją bowiem obiekty we wszechświecie, nazywane *pulsarami*, które traktujemy jako

wysoce precyzyjne zegary. Są to wirujące gwiazdy neutronowe o silnym polu magnetycznym, które niczym latarnia morska emitują pulsujące sygnały radiowe w bardzo krótkich odstępach czasu. Przesunięcia czasu przybycia sygnału wynoszą mniej niż 100 nanosekund (nanosekunda to jedna miliardowa sekundy). Najdokładniej wymierzonym systemem jest odkryta w 1974 roku gwiazda podwójna PSR 1913+16 (liczby te oznaczają jej położenie na niebie), zawierająca pulsar. Okres orbitalny tego *układu podwójnego* zmniejszył się od momentu jego odkrycia o ponad 35 sekund, co w pełni pokrywa się z teorią Einsteina. Jedynym wyjaśnieniem tego efektu jest emisja fal grawitacyjnych. Bardzo szczęśliwym trafem okazało się odkrycie wiosną 2003 roku *pulsara podwójnego*, czyli układu gwiazd podwójnych, w którym *obie* gwiazdy są pulsarami. W przypadku takiego układu ogólna teoria względności prognozuje kurczenie się promienia orbity o siedem milimetrów dziennie i zmniejszenie okresu orbitalnego o 100 nanosekund dziennie, co można było potwierdzić z dokładnością do jednego procenta. W obiektach tych można również dostrzec opisane w ostatnim podrozdziale zjawiska znane z Układu Słonecznego, takie jak choćby grawitacyjna dylatacja czasu czy precesja peryhelium (w tym kontekście określana jako obrót peryastronu). Zamiast 43 sekund kątowych na 100 lat, jak w przypadku Merkurego, przesunięcie osiąga tu aż 17° rocznie, co odpowiada jednemu obrotowi peryastronu w ciągu około 21 lat. Triumf Einsteina nie mógł być większy!

Czarne dziury

Przekonaliśmy się, że dla teorii względności w Układzie Słonecznym istnieje dokładne rozwiązanie równań Einsteina w formie czasoprzestrzeni Schwarzschilda (rys. 1.10). Opisuje ona zależności geometryczne poza Słońcem. Można udowodnić, że mamy do czynienia z jedynym statycznym rozwiązaniem tych równań. Co się stanie, gdy wyobrazimy sobie stopniowe skracanie promienia Słońca bez jednoczesnego ubytku jego masy? Powstała w ten sposób czasoprzestrzeń opisuje jeden z najbardziej fascynujących obiektów – *czarną dziurę*. Pojęcie

wprowadzone przez Johna Wheelera w 1967 roku wyraża najistotniejszą właściwość tego tworu: jego pole grawitacyjne tak silne, że nie wymknie się z niego nawet światło, czyli żadna informacja. Efekt odchylenia światła jest tak silny, że nawet wyemitowane na zewnątrz promienie świetlne zostaną przechwycone wskutek działania grawitacji. W przypadku masy Słońca taka czarna dziura mogłaby powstać, gdyby promień Słońca skurczył się do wartości poniżej trzech kilometrów, czyli do wartości znikomej w porównaniu z rzeczywistą długością promienia Słońca, wynoszącego około 700 000 kilometrów. Ta krytyczna wartość nazywana jest także *promieniem Schwarzschilda*. Jest on bezpośrednio proporcjonalny do masy i zależy ponadto tylko od stałej grawitacyjnej i prędkości światła. W przypadku gdy masa spadnie poniżej wartości promienia Schwarzschilda, utworzy się w tym miejscu tak zwany *horyzont zdarzeń*, który oddziela obszar wewnętrzny (właściwą czarną dziurę) od obszaru zewnętrznego. To właśnie z tego obszaru wewnętrznego żadna informacja nie może wydostać się na zewnątrz. Zgodnie z geometrią czasoprzestrzeni można też powiedzieć, że w przypadku obserwatorów znajdujących się na zewnątrz ten wewnętrzny obszar nigdy nie stanie się przeszłością. W związku z tym część czasoprzestrzeni jest przyczynowo całkowicie odłączona od reszty świata, a to naprawdę osobliwe wyobrażenie. Na rysunku 1.18 przedstawiono kolaps gwiazdy i narodziny czarnej dziury. Zależności przyczynowe są zaznaczone za pomocą kilku lokalnych stożków świetlnych. Można zauważyć, że stożki wewnątrz horyzontu zdarzeń są skierowane do wewnątrz, tak aby wszelkie możliwe sygnały, które mogą mnożyć się na stożku świetlnym lub w jego wnętrzu, musiały biec do wewnątrz. Sam horyzont jest styczny do lokalnych stożków świetlnych i można sobie wyobrazić, że jest zbudowany z promieni świetlnych.



1.18. Diagram czasoprzestrzeni zapadającej się gwiazdy, z której powstaje czarna dziura (jeden wymiar przestrzeni został pominięty)

Co znajduje się wewnątrz horyzontu zdarzeń? Równania Einsteina dopuszczają tylko jedną możliwość: masa zostaje ściśnięta do jednego punktu, tworząc obszar nieskończenie wielkiej czasoprzestrzennej krzywizny. W takim przypadku mówi się o tak zwanej *osobliwości*. Z uwagi na to, że z czarnej dziury nie może wydostać się na zewnątrz żadna informacja, nie możemy także dostrzec osobliwości. W istnienie takiej osobliwości rzeczywiście trudno jest uwierzyć i przypuszczalnie nie mieści się ona nawet w ramach teorii, która jest bardziej fundamentalna niż teoria Einsteina. Więcej na ten temat w następnych rozdziałach.

Nie tylko odchylenie światła w czarnej dziurze jest ogromne. To samo odnosi się do wszystkich innych zjawisk teorii względności. Podczas gdy w Układzie Słonecznym są one zaledwie małymi korektami fizyki Newtona, w czarnej dziurze determinują bieg zdarzeń. Za przykład może posłużyć grawitacyjne opóźnienie w czasie, które w obrębie horyzontu zdarzeń staje się nieskończenie wielkie. Świadczy to o tym, że w odniesieniu do bardzo odległych obserwatorów zegary zbliżające się do horyzontu zdarzeń stopniowo zwalniają i ostatecznie zatrzymują swój bieg. Sytuacja ta dotyczy w równym stopniu wszystkich procesów, co oznacza, że z perspektywy zewnętrznego obserwatora na horyzoncie zdarzeń zamiera każdy ruch. Nigdy nie uda się zaobserwować przeniknięcia czegokolwiek do czarnej dziury (porównaj z rys. 1.18). Obserwator, który wpadnie do czarnej dziury, nie zauważy natomiast żadnej zmiany w ruchu swojego zegarka. Wynika to bezpośrednio z zasady równoważności. W skończonym czasie wpadnie niechybnie w osobliwość.

Sygnał, wyemitowany na zewnątrz tuż poza horyzontem zdarzeń, doznaje dodatkowo znacznego przesunięcia ku czerwieni, czyli efektu, który poznaliśmy już w związku z zasadą równoważności. Przesunięcie ku czerwieni powoduje, że obiekt wpadający do czarnej dziury staje się bardzo szybko niewidzialny, nawet jeśli dla zewnętrznego obserwatora zdaje się on zastygać na krawędzi horyzontu zdarzeń.

Wyobraziliśmy sobie przemianę Słońce w czarną dziurę poprzez stopniowe zmniejszanie jego promienia i zachowanie stałej masy. Prawdziwe Słońce nie zmieni się jednak w czarną dziurę. Po wyczerpaniu się zapasów energii jądrowej będzie ono nadal trwało w marnym bycie jako tak zwany biały karzeł – gwiazda wielkości Ziemi. Czarne dziury powstają natomiast z bardzo masywnych gwiazd, które zapadają się po zakończonej syntezie jądrowej. Pozostała po kolapsie masa musi równać się co najmniej trzykrotnej masie Słońca. To, że w takich przypadkach powstanie czarnej dziury w wyniku kolapsu jest nieuniknione, wynika z bardzo ogólnych przyczyn, które mogą przyjąć postać twierdzeń matematycznych. Takie „twierdzenia o osobliwości” zostały udowodnione

przez dwóch brytyjskich fizyków, Rogera Penrose'a i Stephena Hawkinga, w latach sześćdziesiątych XX wieku. Z powodu równoważności masy i energii każda forma energii wytwarza pole grawitacyjne. Gdy postęp kolapsu jest już dostatecznie zaawansowany, uwalniana energia samodzielnie wytwarza przyciągające pole grawitacyjne, które ponownie przyspiesza kolaps, i tak dzieje się aż do osiągnięcia stanu osobliwości.

O ile proces powstawania czarnych dziur jest skomplikowany, o tyle prosty jest jego rezultat. John Wheeler wyraził to w oryginalny sposób, mówiąc, że czarne dziury nie mają włosów, a mianowicie, gdy czarne dziury osiągają swój stan końcowy, można je opisać za pomocą trzech parametrów: masy, momentu pędu i ładunku elektrycznego. Naładowane czarne dziury nie odgrywają w astrofizyce żadnej roli, a zatem można ograniczyć się do dwóch wielkości: masy i momentu pędu. Wszystkie inne właściwości („włosy") zostają wcześniej oddane, na przykład w formie fal grawitacyjnych. Jeżeli dwie dziury mają zgodne parametry, to nie istnieje już nic, co mogłoby je od siebie odróżnić. Nie lada różnica w porównaniu z różnorodną naturą zwykłych gwiazd!

Dokładne rozwiązanie równań Einsteina, opisujące wirującą czarną dziurę, zostało odkryte przez nowozelandzkiego naukowca Roya Kerra w 1963 roku. Określają je dwa parametry: masa i moment pędu. W odróżnieniu od rozwiązania Schwarzschilda nie jest już sferycznie symetryczne, wykazuje jednak symetrię osiową (symetrię obrotową), która odnosi się do osi obrotu.

Czarne dziury to nie tylko formalne rozwiązania skomplikowanych równań na kartce papieru. Można je zaobserwować również w przyrodzie. Z uwagi na to, że same pozostają niewidoczne, rozpoznaje się je po wpływie, jaki wywierają na otoczenie. Czarne dziury towarzyszą układom gwiazd podwójnych, zasysając w charakterystyczny dla siebie sposób gaz ze zwykłej gwiazdy. Odkryto już wiele takich układów gwiazd. Dodatkowo w centrum prawie każdej galaktyki czyha czarna dziura o gigantycznej masie. W centrum naszej Drogi Mlecznej, znajdującym się w odległości 26 000 lat świetlnych od nas, istnieje czarna dziura wielkości

ponad trzech milionów mas Słońca. Można ją zaobserwować na podstawie wpływu, jaki wywiera na ruchy okrążających ją gwiazd w jej otoczeniu. Prawdopodobnie, tak jak wszystkie czarne dziury występujące w przyrodzie, posiada ona moment pędu, a zatem można ją opisać dzięki rozwiązaniu Kerra.

Czarne dziury mają horyzont zdarzeń. Pytanie, czy osobliwości mogą powstawać także bez horyzontu, należy do najciekawszych otwartych pytań w teorii Einsteina. Taka osobliwość byłaby „naga", czyli widzialna z zewnątrz. Jej istnienie oznaczałoby dla teorii wielkie tarapaty, ponieważ do naszego wszechświata mogłyby docierać z niej dowolne rzeczy, których pojawienia się teoria nie byłaby w stanie przewidzieć. W związku z tym w 1969 roku Penrose postulował hipotezę „kosmicznej cenzury", która zabrania istnienia takim nagim osobliwościom. Ów postulat nie oznacza automatycznie powstania odpowiedniej zasady. Czy taka cenzura rzeczywiście pojawia się w realistycznej, czyli występującej w przyrodzie sytuacji, nie zostało dotychczas ani udowodnione, ani obalone.

Ile wymiarów ma świat?

Przestrzeń ma trzy wymiary i łącznie z czasem tworzy czterowymiarową czasoprzestrzeń. Dlaczego ktoś miałby w to wątpić? W 1919 roku matematyk z Królewca Theodor Kaluza (1885–1954) zadał sobie pytanie, czy grawitację i elektromagnetyzm można opisać za pomocą jednej spójnej teorii. Podejście redukcjonistyczne, czyli opis różnych sił w ramach jednej teorii, miało przecież odnieść sukces w fizyce. Kaluza wpadł na genialny pomysł, aby oba oddziaływania zapisać za pomocą jednej teorii geometrycznej. W tym celu konieczne było jednak podjęcie dość kontrowersyjnej decyzji i powołanie się na nowy, *piąty* wymiar świata. Kaluza przyjrzał się równaniom pola Einsteina w pięciu wymiarach czasoprzestrzeni i stwierdził, że część pięciowymiarowej metryki odpowiada normalnej czterowymiarowej metryce pochodzącej z teorii względności Einsteina, a pozostała część może zostać powiązana z polem elektromagnetycznym. Światło nie byłoby wtedy niczym innym

niż manifestacją piątego wymiaru. Z uwagi na to, że na co dzień nie zauważamy obecności dodatkowego wymiaru, Kaluza sformułował dodatkowy postulat mówiący o tym, iż czterowymiarowe wielkości geometryczne nie mogą zależeć od współrzędnych w kierunku piątego wymiaru.

Kilka lat później szwedzki fizyk Oskar Klein (1894–1977) uzupełnił teorię Kaluzy, uwzględniając w rozważaniach teorię kwantową. W prosty sposób zastosował on również dodatkowy postulat Kaluzy, doczepiając piąty wymiar niczym niewielkie kółko do każdego punktu czterowymiarowej czasoprzestrzeni. Ze względu na swój znikomy rozmiar ten niejako zwinięty piąty wymiar nie byłby zatem bezpośrednio widzialny w naszym makroskopowym świecie. Aby zobrazować tę teorię, można posłużyć się porównaniem do węża ogrodowego, który z dużej odległości wydaje się jednowymiarowy, z bliska jednak wykazuje istnienie drugiego okrągłego wymiaru. Tak zwinięte wyższe wymiary określa się też mianem „skompaktyfikowanych", natomiast teorie zawierające wyższe wymiary nazywa się teoriami Kaluzy-Kleina.

Wyobrażenie o wyższych wymiarach w fizyce pozostawało w cieniu przez dziesiątki lat. Dopiero w latach siedemdziesiątych XX wieku uznano, że znane oddziaływania (do których należały wtedy nie tylko grawitacja i elektromagnetyzm) można w spójny sposób opisać za pomocą jednej teorii o większej liczbie wymiarów czasu i przestrzeni. Teorie Kaluzy-Kleina i ich nowoczesne wersje należą od tej pory do standardowego repertuaru fizyki teoretycznej. Na tym gruncie powstał na przykład obraz czasoprzestrzeni z dziesięcioma lub nawet jedenastoma wymiarami, opisywany w teorii strun w części 5 niniejszej książki.

Wprawdzie jak dotąd żaden eksperyment nie dowiódł istnienia wyższych wymiarów, jednak wyobrażenie świata, który w rzeczywistości posiada więcej niż trzy wymiary, nie stało się w związku z tym mniej atrakcyjne. Nawet Einstein pisał do Kaluzy 9 grudnia 1921 roku, że jego myśl jest ujmująca i musi w niej być coś prawdziwego.

2. Teoria kwantowa

Jeśli istnieje poczucie rzeczywistości, a nikt nie podważy jego prawa do istnienia, to musi istnieć też coś, co można by nazwać poczuciem ewentualności.

Robert Musil,
Człowiek bez właściwości

Co opisuje teoria kwantowa?

Początek nowej epoki wiąże się zazwyczaj z jakąś znaczącą datą, tak jak na przykład zajęcie Bastylii zapoczątkowało zmiany w życiu politycznym. W fizyce jest podobnie. Za początek stulecia w dziedzinie elektrodynamiki można uznać podróż Alessandra Volty do Paryża w 1801 roku, gdzie w obecności Napoleona z powodzeniem zaprezentował on swoje eksperymenty z elektrycznością.

Narodziny teorii kwantowej wiążą się z wykładem Maxa Plancka (1858–1947) wygłoszonym 14 grudnia 1900 roku przed Niemieckim Towarzystwem Fizycznym. To wtedy usłyszano o wielkości, która miała zrewolucjonizować fizykę jak żadna inna do tej pory, zwanej później kwantem działania Plancka *h*. Planck wysunął z pozoru niewinne założenie, że atomy mogą przyjmować i oddawać energię promieniowania elektromagnetycznego tylko w całkowitych wielokrotnościach energii podstawowej. Wielkość tej energii podstawowej jest iloczynem *h* i częstotliwości promieniowania. Planck pisał później, że było to czysto formalne założenie i niewiele sobie po nim obiecywał. Wspomniane występowanie energii w „pakietach”, czyli kwantach, nadało nazwę całej teorii.

Planck przyjął takie formalne założenie, aby opisać eksperymenty, które do tej pory uparcie umykały próbom każdego teoretycznego

wyjaśnienia – eksperymenty dotyczące promieniowania ciał doskonale czarnych. Ciała doskonale czarne to opisane najpierw przez fizyka Gustava Roberta Kirchhoffa (1824–1887) ciała wyidealizowane, których zdolność absorpcyjna posiada maksymalną możliwą wartość. Gdy ciało takie otacza pustkę, ściany znajdują się w równowadze z wyemitowanym przez to ciało promieniowaniem, które tkwi wewnątrz. Energia promieniowania jest przy tym uzależniona od temperatury ciała. Gdy w ścianie zostanie wywiercona niewielka dziura, promieniowanie może wydostać się z wnętrza pustki na zewnątrz i zostać poddane eksperymentalnym badaniom. Z tego powodu nazywane jest ono także „promieniowaniem pustki". Jednocześnie okazało się, że zmierzone widmo promieniowania (pod pojęciem widma promieniowania rozumiany jest rozkład energii promieniowania poprzez częstotliwość promieniowania) nie może być wyjaśnione na podstawie teorii elektrodynamiki Maxwella (zob. część 1). Teoria Maxwella zdawała się przewidywać nawet nieskończenie wielką energię dla wysokich częstotliwości (w „nadfiolecie"), co określano jako katastrofę w nadfiolecie. Założenie *ad hoc* Plancka o pakietach energii zapobiegło tej katastrofie, nie dostarczyło jednak początkowo żadnego głębszego wyjaśnienia.

W istocie poszukiwania spójnej teorii fizycznej trwały do 1927 roku, gdy założenia formalizmu teorii kwantowej (mechaniki kwantowej) były już w znacznej mierze opracowane. Do pionierów w tej dziedzinie, ale nie do jej twórców, należeli obok Plancka przede wszystkim Einstein, Niels Bohr (1885–1962) i Louis de Broglie (1892–1987). Einsteinowi udało się w swoim *annus mirabilis*, to jest w 1905 roku, wyjaśnić efekt podstawowy, czyli efekt fotoelektryczny. Chodzi tu o uwolnienie elektronów z metalu wskutek naświetlania. Energia tych elektronów jest zależna wyłącznie od częstotliwości światła, a nie od jego natężenia – w przeciwieństwie do oczekiwań związanych z teorią Maxwella. Einstein wyjaśnił taki wynik dzięki swojej „hipotezie kwantów światła". Nie tylko atomy w ciele, jak założył Planck, mogą pobierać i oddawać energię wyłącznie w pakietach, również samo światło składa się z dyskretnych pakietów energii, kwantów światła lub *fotonów*. Ich energia ma być ponownie proporcjonalna do

częstotliwości promieniowania, z kwantem działania *h* jako współczynnikiem proporcjonalności. W związku z tym energia uwolnionych elektronów zależy od częstotliwości światła. Wprowadzenie kwantów światła było najbardziej rewolucyjnym dokonaniem Einsteina. Początkowo założenie to zostało zdecydowanie odrzucone przez większość fizyków z Planckiem na czele. Niemniej to właśnie ono, a nie teoria względności, przyniosło później Einsteinowi Nagrodę Nobla. Obecnie zastosowanie efektu fotoelektrycznego, na przykład w ogniwach słonecznych lub aparatach cyfrowych, należy do codzienności. Istnienie kwantów światła uwidoczniło fakt, że teoria Maxwella nie mogła być teorią uniwersalną.

Zjawiska dyskretne zauważono też w innym miejscu. Atomy, jak na przykład atom wodoru, emitują promieniowanie jedynie w określonych częstotliwościach dyskretnych. W 1913 roku Niels Bohr za pomocą hipotezy *ad hoc* poprawnie opisał widmo wodoru. Zgodnie z nią elektron krążący wokół jądra porusza się tylko po określonych dyskretnych orbitach, a wyższym energiom odpowiadają dłuższe promienie. Podczas przejścia na niższą orbitę emitowana jest określona energia w formie kwantu światła, który odpowiada konkretnej częstotliwości, a tym samym także określonej linii w widmie wodoru. W szczególności powinien istnieć stabilny stan podstawowy, czyli minimalny promień, poniżej którego elektron nie może zejść. To, że atomy składają się z ujemnie naładowanych elektronów krążących wokół dodatnio naładowanych jąder, zostało odkryte dopiero dwa lata wcześniej w wyniku słynnych eksperymentów przeprowadzonych przez Ernesta Rutherforda (1871–1937). W 1911 roku Rutherford wystrzeliwał cząstki alfa (jądra helu) na złote folie i zauważył przy tym, że znajdujący się w atomach dodatni ładunek elektryczny koncentruje się w stosunkowo niewielkich jądrach atomu i nie jest rozmieszczony równomiernie w całym atomie, jak uważano wcześniej (analogicznie do ciasta, w którym elektrony są zanurzone jak rodzynki).

W myśl teorii Maxwella elektron krążący wokół maleńkiego jądra musiałby w krótkim czasie uderzyć w jądro i stabilna materia nie mogłaby

istnieć. Również w tym przypadku okazało się, że elektrodynamika klasyczna nie jest uniwersalna. Model atomu Bohra postulował tylko brak zderzenia z jądrem, ale jeszcze tego nie wyjaśniał.

Światło, które teoria Maxwella opisuje jako falę, posiada więc także cechy cząstki. Czy zatem cząstki, podobnie jak elektron, nie powinny mieć czasem cech fali? W 1923 roku taką hipotezę odważył się wysunąć Louis de Broglie w swojej pracy doktorskiej. Według niego każda cząstka posiada też długość fali, nazwaną później „długością fali de Broglie'a". Długość fali zależy więc od kwantu działania i jest tym większa, im mniejszy jest pęd cząstki. Jeżeli wszystkie cząstki również zachowują się jak fale, to jak brzmi odpowiednie dla tego stanu rzeczy równanie falowe? Tego wyzwania podjął się z kolei Erwin Schrödinger (1887–1961), po tym jak w 1925 roku na Uniwersytecie Zuryskim mówił o pracach de Broglie'a. Zainspirowało go to i podczas bożonarodzeniowego urlopu na przełomie 1925 i 1926 roku w szwajcarskiej górskiej wiosce Arosa udało mu się opracować słynne równanie Schrödingera, jedno z najważniejszych równań ostatniego stulecia. Chodzi o równanie falowe dla funkcji falowej Ψ (psi), które ma szerokie zastosowanie i jest podstawowym równaniem mechaniki kwantowej. Już w 1927 roku udało się potwierdzić eksperymentalnie falowy charakter elektronów. Dzięki swojemu równaniu Schrödinger stał się jednym z twórców teorii kwantowej. Z równania Schrödingera dyskretne wartości energii widm wynikają automatycznie, gdy zostaną ustalone warunki brzegowe dla funkcji falowej Ψ.

Obiekty mikroskopowe mogą zatem zachowywać się czasem jak cząstki, a czasem jak fale, przy czym „cząstka" odnosi się tu do sytuacji, w których obiekt został zlokalizowany. Poniżej przekonamy się, jak można zrozumieć poszczególne zachowania na podstawie funkcji falowej Schrödingera. Niezależnie od ich zachowania elementarne mikroobiekty, takie jak elektron czy proton, często określa się mianem cząstek, a odpowiednią dziedzinę fizyki nazywa się fizyką cząstek elementarnych, którym to określeniem także będziemy się posługiwać.

W rzeczywistości formalizm mechaniki kwantowanej został opracowany już kilka miesięcy przed pracami Schrödingera, lecz w dużo mniej przejrzystej odsłonie. Werner Heisenberg (1901–1976), Max Born (1882–1970) i Pascual Jordan (1902–1980) sformułowali w Getyndze tak zwaną mechanikę macierzową, która miała stać się później odpowiednikiem teorii Schrödingera. Za pomocą mechaniki macierzowej Wolfgang Pauli (1900–1958) dokładnie obliczył widmo atomu wodoru jeszcze przed Schrödingerem, choć przy użyciu równania Schrödingera obliczenia są znacznie prostsze. Wraz z ukończeniem teorii kwantowej w fizyce rozpoczęła się nowa epoka, co też natychmiast odczuli wszyscy zainteresowani.

W październiku 1927 roku w Brukseli odbył się jeden ze słynnych kongresów Solvaya, nazwanych tak od imienia ich inicjatora, belgijskiego chemika Ernesta Solvaya. W kongresie tym uczestniczyli niemal wszyscy współtwórcy teorii kwantowej. Doszli oni do wniosku, że formalizm został już opracowany w dostatecznym kształcie i można skupić się na jego interesujących zastosowaniach w fizyce atomowej. Wydarzenie to zapoczątkowało trwającą właściwie do tej pory debatę na temat znaczenia i interpretacji teorii kwantowej. Jak na przykład należy rozumieć fakt, że naładowana cząstka w komorze mgłowej (komorze wypełnionej gazem) ukazuje się jako kreska, choć funkcja falowa, do której dąży cząstka, jest sferycznie symetryczna, czyli zdaje się nie mieć nic wspólnego ze śladem cząstki? Wrócimy jeszcze do tego później.

W tamtym czasie uważano, że teoria kwantowa jest teorią fizyczną, która opisuje atomy i ich oddziaływanie z promieniowaniem elektromagnetycznym, lecz nie można jej zastosować w przypadku jeszcze mniejszych wymiarów (jak choćby jąder atomów, składających się z protonów i neutronów). W kolejnych dziesięcioleciach okazało się jednak, że teoria kwantowa doskonale radzi sobie z opisem wszystkich obiektów i ich oddziaływań w obszarze mikroskopowym i poza nim. W związku z tym często mówi się o *uniwersalizmie* teorii kwantowej. Tworzy ona teoretyczne ramy, które zdają się opisywać dosłownie

wszystko z wyjątkiem grawitacji, stojącej jak dotąd na uboczu. Jak mogliśmy się przekonać w części 1, grawitacja została z powodzeniem opisana przez ogólną teorię względności Einsteina, którą określa się jako teorię klasyczną (wszystkie teorie w fizyce poza teorią kwantową określa się mianem klasycznych). Poniżej zobaczymy, że istnieją uzasadnione powody, aby również grawitację opisać za pomocą teorii kwantowej, a zatem teorii grawitacji kwantowej, co może mieć istotny wpływ na nasze rozumienie wszechświata.

Teoria kwantowa przenika do naszego życia codziennego dzięki swoim technicznym zastosowaniom. Na niej opiera się sposób funkcjonowania lasera wykorzystywanego na przykład w płytach kompaktowych (CD) lub okulistyce. Bez lasera nie można się już obyć zwłaszcza w medycynie. Dzięki spektroskopii magnetycznego rezonansu jądrowego udaje się uzyskać obraz wnętrza człowieka, a terapia protonowa ma zastosowanie w leczeniu nowotworów. Innym przykładem jej użycia są półprzewodniki w chipach komputerowych. W 2007 roku dwóm naukowcom, Peterowi Grünbergowi i Albertowi Fertowi, przyznano Nagrodę Nobla w dziedzinie fizyki za odkrycie „efektu kwantowego” (tzw. gigantycznego magnetooporu), który umożliwił znaczne zwiększenie pojemności dysków twardych w naszych komputerach.

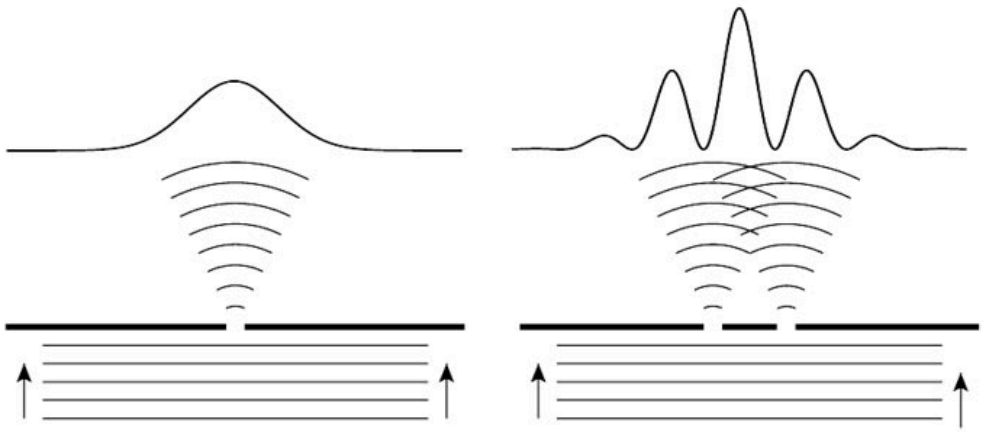
Zastosowania tej teorii nie są jednak tematem niniejszej książki. Nasza uwaga skupia się raczej na tym, aby przedstawić rolę teorii kwantowej w celu lepszego zrozumienia przyrody i drogi do jej wszechstronnego opisu. Z uwagi na to, że fizyka jest podstawą innych nauk przyrodniczych, na teorii kwantowej bazują przecież także chemia i biologia.

Stany splątane i obalenie rzeczywistości lokalnej

Podstawą teorii kwantowej jest zasada superpozycji (zasada nakładania się). Zasada ta jest znana również z fizyki klasycznej, jak choćby z teorii nakładania się fal. Została ona zobrazowana na przykładzie fali wodnej, która płynie równocześnie przez dwa kanały i łączy się poza nimi. Gdy grzbiet fali napotka dolinę fali, znoszą się wzajemnie i następuje ich

wygaszenie. Jeżeli jednak grzbiet fali napotka grzbiet, a dolina zetknie się z doliną, to mamy do czynienia ze wzmocnieniem. Grzbiety bądź doliny sumują się. Wystąpienie wzmocnienia lub wygaszenia zależy od przestrzennej odległości między falami składowymi, określanej mianem różnicy dróg.

Zjawiska wygaszenia i wzmocnienia zwane są interferencjami. Istnieją one nie tylko w odniesieniu do fal wodnych. W 1801 roku brytyjski fizyk Thomas Young (1773–1829) jako pierwszy stwierdził występowanie tych zjawisk w przypadku światła. Panowało wtedy (choć nie wyłącznie) przekonanie zgodne z teoriami Newtona, że światło składa się z maleńkich cząstek. Eksperymenty Younga wpłynęły na zmianę tej opinii i doprowadziły do powstania falowej teorii światła. Dopiero teoria kwantowa uwidoczniła, że światło nie jest ani falą, ani cząstką, lecz czymś zupełnie innym, o wiele bardziej niejasnym, co czasem przybiera formę fali, a czasem cząstki.



2.1. Eksperyment podwójnej szczeliny Thomasa Younga

Na rysunku 2.1 przedstawiono schemat eksperymentu interferencyjnego Younga z podwójną szczeliną podświetlaną od dołu. Po lewej stronie rysunku jedna ze szczelin jest zamknięta, przez co widać tylko jedną z nich. Po prawej stronie rysunku obie szczeliny są otwarte i na ekranie można zaobserwować rozkład jasności, która nie jest sumą obrazów obu

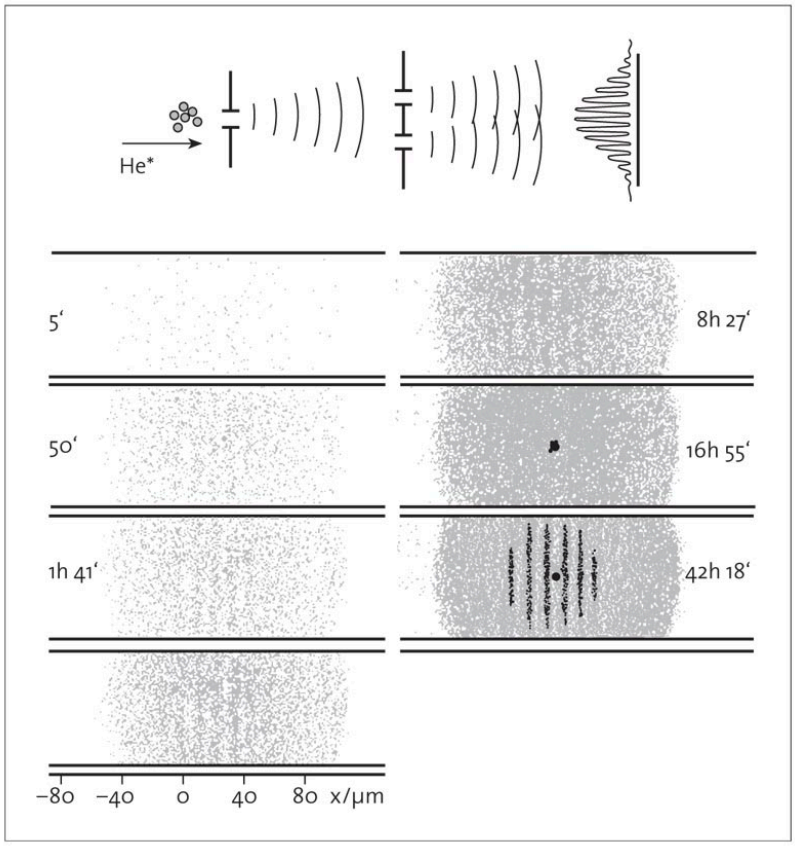
pojedynczych szczelin, lecz składa się z kolejnych jasnych i ciemnych prążków. Jest to wzorzec interferencyjny. Jasne prążki oznaczają wzmocnienie, a ciemne wygaszenie fal. Odstępy między prążkami są proporcjonalne do długości fal świetlnych i odwrotnie proporcjonalne do odległości między szczelinami. Za czasów Younga nie było jednak jasne, co jest przyczyną takich drgań. Dopiero znacznie później odkryto, że chodzi o drgania pola elektromagnetycznego, opisane przez równania Maxwella (część 1). Okres drgań w optyce wynosi mniej niż jedną femtosekundę (to jest 10^{-15} sekundy), dlatego jest niewidoczny dla naszego oka, które może dostrzec jedynie średnie natężenie światła.

Zasada superpozycji w teorii kwantowej jest uniwersalna i dotyczy dowolnych stanów kwantowych. Przyczyną tego faktu jest to, że funkcja falowa Schrödingera opisuje wszystkie obiekty mikroskopowe, choć jest zdefiniowana jedynie dla szczególnego przypadku cząstki w normalnej trójwymiarowej przestrzeni. Przestrzeń służąca za podstawę jest przestrzenią abstrakcyjną, czyli tak zwaną przestrzenią konfiguracyjną systemu, w którym każda cząstka posiada trzy wymiary, co w przypadku dwóch cząstek oznacza sześciowymiarową przestrzeń, w przypadku trzech – dziewięciowymiarową i tak dalej. W rzeczywistych sytuacjach przestrzeń konfiguracyjna jest zatem wielowymiarowa. Na tym polega najistotniejsza różnica w stosunku do fal klasycznych, które rozprzestrzeniają się zawsze w trzech wymiarach przestrzeni. Dalej przekonamy się również, że wspomniany już charakter cząstek obiektów kwantowych ma swoje źródło w tej wielowymiarowej funkcji falowej.

Stan kwantowy powstający w przypadku superpozycji jest całkowicie nowym stanem, a nie jedynie mieszanką poszczególnych składowych. Mogą wtedy powstać całkowicie nowe rodzaje cząstek. Jeżeli neutralny mezon K znajdzie się w superpozycji ze swoją antycząstką, powstanie inny mezon K o nowych właściwościach. Podobnie jest u słabo oddziałujących neutrin. Natomiast w przypadku światła laserowego dochodzi do superpozycji stanów kwantowych z różnymi liczbami cząstek.

W myśl teorii kwantowej zjawiska interferencji mogą wystąpić w przypadku elektronów, atomów, molekuł i wielu innych obiektów. Eksperyment podwójnej szczeliny z elektronami jest imponujący ze względu na fundamentalne znaczenie elektronów. Po raz pierwszy został on przeprowadzony w 1959 roku przez Clausa Jönssona i Gottfrieda Möllenstedta na Uniwersytecie Eberharda i Karola w Tybindze (eksperyment uznany przez brytyjskie czasopismo „Physics World" w 2002 roku za najpiękniejszy fizyczny eksperyment wszech czasów). Sytuacja w tym eksperymencie jest analogiczna do rysunku 2.1, tylko zamiast światła emitowane są właśnie elektrony. Współczesna wersja tego eksperymentu została przeprowadzona w 2005 roku w Garching koło Monachium przy współudziale Instytutu Optyki Kwantowej im. Maxa Plancka. W ramach eksperymentu dochodzi do dyfrakcji elektronów nie przez szczelinę w przestrzeni, lecz w czasie przy użyciu fali laserowej. Również w tym przypadku mamy do czynienia z obrazem interferencyjnym, przewidzianym przez teorię kwantową.

Współczesne eksperymenty można zorganizować w taki sposób, aby w określonym czasie w ruchu znajdował się zawsze tylko jeden obiekt (jedna „cząstka"). Obiekt taki przechodzi równocześnie przez obie szczeliny, a gdy natrafi na ekran znajdujący się poza szczelinami, pozostawia na nim punkt. Z wielu takich punktów powstaje wzorzec interferencyjny (zob. rys. 2.2).



2.2. Eksperyment podwójnej szczeliny z atomami helu. Można dostrzec, jak z biegiem czasu z poszczególnych zdarzeń powstaje wzorzec interferencyjny.

W ostatnim czasie udało się także przeprowadzić eksperymenty interferencyjne z molekułami. Grupa naukowców, skupiona wokół Antona Zeilingera i Markusa Arndta na Uniwersytecie Wiedeńskim, doprowadziła do interferencji fulerenów C60, czyli molekuł węgla z sześćdziesięcioma lub siedemdziesięcioma atomami, jak i biomolekuł, z których najcięższy składa się ze 108 atomów. Również w tym przypadku za każdym razem tylko jedna molekuła przechodzi równocześnie przez obie szczeliny i interferuje ze sobą. Na fundamentalnym poziomie obraz molekuł znany z podręczników chemii, czyli kuleczek połączonych pałeczkami, nie może być zatem poprawny. W dalszej części dowiemy się, w jakich sytuacjach molekuły przyjmują taką postać. Grupa wiedeńska rozważa nawet przeprowadzenie eksperymentu interferencyjnego z wirusem.

Eksperymenty interferencyjne można przeprowadzać nawet z kondensatem Bosego-Einsteina. Są to układy atomów w takim samym stanie kwantowym, opisane przez jedną funkcję falową, mimo że w takim kondensacie może znajdować się do 10 milionów atomów. Do tego typu eksperymentów niezbędne jest jednak mocne chłodzenie, ponieważ stan kondensatu tworzy się wyłącznie w temperaturze lekko przekraczającej zero absolutne. W innym eksperymencie dochodzi do powstania superpozycji strumienia nadprzewodzących (tzn. poruszających się z zerową rezystancją) elektronów z takim samym strumieniem płynącym w przeciwnym kierunku. Również w tym przypadku występują zjawiska interferencji, przepowiadane przez teorię kwantową. Sytuację tę należy jednak wyraźnie odróżnić od klasycznego nakładania się na siebie dwóch stanów. W przypadku wystąpienia superpozycji dwóch klasycznych strumieni płynących w przeciwnych kierunkach nastąpi całkowite ich zniesienie, czyli nie dojdzie do niczego szczególnego.

Jak widać na rysunku 2.2, wzorzec interferencyjny powstaje na ekranie dopiero w wyniku nagromadzenia się bardzo wielu pojedynczych punktów. W teorii kwantowej względna częstość tych punktów interpretowana jest jako prawdopodobieństwo możliwości zaobserwowania jednego punktu w określonym miejscu na ekranie. Interpretacja ta wynika z fundamentalnej pracy Maxa Borna z 1926 roku. Według niej to kwadrat funkcji falowej Ψ podaje prawdopodobieństwo możliwości zaobserwowania odpowiedniego obiektu (np. elektronu) w określonym miejscu. Ta interpretacja prawdopodobieństwa teorii kwantowej świetnie sprawdziła się w eksperymentach i z tego powodu Ψ nazywana jest często „falą prawdopodobieństwa”. Czy zatem w fizyce na płaszczyźnie fundamentalnej istnieją jedynie prawdopodobieństwa? W tym miejscu wskazywano często na indeterministyczny (nieokreślony) charakter praw przyrody, wynikających z teorii kwantowej, co nie spotkało się z entuzjazmem ze strony wszystkich fizyków. Nawet Einstein stale podkreślał, że nie wierzy, aby Bóg grał w kości. Niemniej jednak należy zaznaczyć, że samo równanie Schrödingera jest deterministyczne, ponieważ funkcja falowa przyjęta w danym czasie jest ważna dla każdego

czasu, zarówno w przyszłości, jak i w przeszłości. Charakter prawdopodobieństwa przejawia się wtedy, gdy w grę wchodzą klasyczne wielkości, takie jak położenie czy pęd cząstki.

Z funkcji falowej wynika, że nie można równocześnie dokładnie określić położenia i pędu cząstki. Wyraża to słynna zasada nieoznaczoności, opublikowana przez Heisenberga w 1927 roku. Granice są ustalane ponownie przez kwant działania. Alternatywne określenie „zasada nieostrości" jest nieco mylące, gdyż sugeruje, że obiekt sam w sobie ma określone położenie i określony pęd. W rzeczywistości w teorii kwantowej nie istnieją już żadne orbity, a jedynie prawdopodobieństwa znalezienia cząstki w danym punkcie. Elektrony w atomach nie krążą więc po orbitach wokół jądra, inaczej niż zakładał to model Bohra, lecz są charakteryzowane przez funkcje falowe, na podstawie których można wyliczyć prawdopodobieństwo znalezienia ich w określonych miejscach. Co więcej, funkcje falowe, które należą do stałej energii elektronu, są statyczne, czyli w ogóle nie zależą od czasu. Ponadto istnieje podstawowy stan najniższej energii, poniżej którego elektron nie może zejść. W ten sposób mechanika kwantowa uniemożliwia klasyczny kolaps do jądra atomu. Tego, w jaki sposób kształtują się orbity cząstek, dowiemy się, przynajmniej w przybliżeniu, w dalszej części tej książki.

Zasada nieoznaczoności między energią a czasem obowiązuje również w mechanice kwantowej, jednak różni się w swojej interpretacji od tej występującej między pędem a położeniem. Wiąże się to z tym, że czas nadal jest parametrem klasycznym – chodzi tu o czas absolutny według teorii Newtona, przejęty w niezmienionej postaci przez mechanikę kwantową. Nieoznaczoność energii i czasu oznacza w gruncie rzeczy, że żywotność niestabilnego stanu kwantowego jest odwrotnie proporcjonalna do nieoznaczoności energii, przy czym stałą proporcjonalności jest ponownie stała Plancka h, a zatem im mniejsza nieoznaczoność energii, tym większa związana z tym stanem żywotność. W przypadku atomów nieoznaczoność energii wyraża się w szerokości linii widmowych.

Obiekty kwantowe wykazują jeszcze więcej niezwykłych cech. Oprócz momentu pędu orbitalnego posiadają one również wewnętrzny moment pędu, czyli spin, który nie ma żadnego odpowiednika w fizyce klasycznej. Jest on proporcjonalny do kwantu działania, przy czym współczynnik proporcjonalności może przyjmować tylko wartości całkowite lub połówkowe. Cząstki o spinie połówkowym nazywa się fermionami, a o spinie całkowitym – bozonami. Do fermionów należą między innymi elektron, proton i neutron, których spin = ½, zaś do bozonów – między innymi foton o spinie = 1. Z uwagi na to, że spin posiada także kierunek, jego składowa w odniesieniu do danej osi dla cząstki o spinie = ½ może być zwrócona w kierunku dodatnim lub ujemnym, czyli wynosić plus lub minus ½ (mowa tu o „spinie w górę" lub „spinie w dół"). Na przykład dla fotonu będą to wartości plus lub minus 1. Mechanika kwantowa zakłada też dla spinu zasadę nieoznaczoności bądź nieostrości. W związku ze zróżnicowanymi osiami przestrzennymi składowych spinu nie można równocześnie dokładnie określić.

Fermiony i bozony różnią się od siebie zasadniczo. Każda para fermionów nie może znajdować się w tym samym stanie, w przeciwieństwie do bozonów. Tak brzmiała sformułowana przez Pauliego zasada wykluczenia, nazywana również zakazem Pauliego, która wynika z formalizmu teorii kwantowej. Z uwagi na to, że obowiązuje dla elektronów, jest podstawą układu okresowego pierwiastków, a przez to całej chemii. Można także wykazać, że materia bez tej zasady byłaby niestabilna, czyli nie mogłoby istnieć życie. Fizyka klasyczna nie pozwala zatem nawet w podstawowym stopniu na wyjaśnienie naszej egzystencji.

W odróżnieniu od klasycznych obiektów, obiekty kwantowe są nierozróżnialne – każdy elektron wygląda tak samo jak inny. Leibniz twierdził nawet, że nie jest możliwe istnienie dwóch nierozróżnialnych przedmiotów. Wniosek taki nasunął mu się na podstawie jego „zasady dostatecznego powodu". Gdyby Bóg stworzył dwa identyczne atomy, nie istniałby powód, aby jeden z nich umieścić tu, a drugi w innym miejscu (zamiast odwrotnie). Aby to zademonstrować, wysłał on podobno całe

swoje dworskie towarzystwo do ogrodu pałacowego w Herrenhausen. Ich zadanie polegało na poszukiwaniu dwóch identycznych liści, co było oczywiście z góry skazane na porażkę. Gdyby goście szukali elektronów, byliby w stanie znaleźć tylko jedną wersję każdego z nich.

Kolejnym niezwykłym zjawiskiem teorii kwantowej jest efekt tunelowy. Klasyczne cząstki nie są w stanie pokonać bariery energetycznej, której wartość jest większa niż ich własna energia. W przypadku cząstek kwantowych jest to jednak możliwe. Z uwagi na to, że są opisane przez funkcję falową, istnieje w ich przypadku małe prawdopodobieństwo pokonania bariery energetycznej – fala może przeniknąć jakby pod nią w tunelu. Po natrafieniu na barierę cząstka może pojawić się przed nią albo za nią. Klasyczna cząstka w przypadku mniejszej energii zawsze zostanie odbita. Wiele ważnych procesów nie mogłoby zaistnieć bez efektu tunelowego. Za przykład może posłużyć fuzja jąder wodoru, w wyniku której powstają jądra helu (reakcja syntezy jądrowej). Źródło energii Słońca, a tym samym nasza egzystencja, opiera się w istotnej części na tym właśnie efekcie.

Poznaliśmy zasadę superpozycji jako zasadniczą część teorii kwantowej. Zasada ta dotyczy także stanów kwantowych w układach złożonych. W tym kontekście często pojawia się pojęcie *splątania*. Może tu chodzić o układ, którego części (na przykład różne cząstki) są od siebie bardzo oddalone w przestrzeni. Tylko w takim szczególnym przypadku, gdy stan całkowity (całkowita funkcja falowa) jest równy iloczynowi stanów pojedynczych, układy cząstek są od siebie niezależne. W ogólnym przypadku splątania stanów kwantowych istnieje tylko jeden stan dla całego układu. Z uwagi na to, że stanu tego nie można opisać jako iloczynu stanów poszczególnych układów częściowych, czyli nie da się go już rozszczepić, określa się go, a tym samym również cały układ, jako *nielokalny* lub *nierozdzielny* (nierozłączny). W takim wypadku nie ma po prostu sensu mówić o stanie kwantowym układu częściowego. To właśnie nielokalność przyprawia wielu naukowców o ból głowy, udowadniając, że całość to coś więcej niż tylko suma pojedynczych części.

W 1935 roku aspekt ten po raz pierwszy jasno zaakcentowali w swojej słynnej pracy Albert Einstein, Borys Podolsky i Nathan Rosen, dlatego nazywany jest on od pierwszych liter nazwisk autorów sytuacją EPR lub problemem EPR. O co w nim właściwie chodzi? Obiektem obserwacji jest atom w lekko zmodyfikowanej formie, emitujący cząstki o spinie = ½, które pędzą w przeciwnych kierunkach. Cząstki znajdują się przy tym w stanie splątanym, w którym pojedynczej cząstce nie jest przypisana żadna wartość spinu. W stanie tym spiny są „antyskorelowane". Gdy pomiar składowej spinu pierwszej cząstki względem kierunku pionowego dostarczy wynik „spin w górę", to w przypadku pomiaru spinu drugiej cząstki wynik musi brzmieć „spin w dół". Gdy dokonuje się natomiast pomiarów spinu w kierunku poziomym i dla pierwszej cząstki wynik wynosi „spin w prawo", to dla drugiej cząstki wynik musi brzmieć „spin w lewo". Antykorelacja ta istnieje w odniesieniu do każdego kierunku.

Dzięki swojej pracy EPR chcieli pokazać, że teoria kwantowa jest niekompletna. Proces myślowy w niej zawarty jest instruktywny. EPR zakładają kryterium lokalności, co znaczy, że w przypadku cząstek, które są od siebie w dostatecznym stopniu oddalone, powinna istnieć możliwość przeprowadzenia pomiaru na jednej z nich bez zakłócania drugiej (i odwrotnie). W pierwszej cząstce można więc zmierzyć składową spinu w kierunku pionowym *lub* poziomym, a w drugiej cząstce pojawi się wartość przeciwna. Z uwagi na to, że pomiar pierwszej cząstki nie zaburza drugiej cząstki, według EPR wszystkie składowe spinu drugiej cząstki muszą mieć ustalone wartości (to samo dotyczy oczywiście także pierwszej cząstki). Mechanika kwantowa nie zezwala jednak na dokładny pomiar składowych spinu w różnych kierunkach równocześnie z powodu zasady nieoznaczoności, dlatego według EPR musi być niekompletna. Powinna zatem istnieć jeszcze bardziej fundamentalna teoria, która pozwoli na obliczenie równoczesnych składowych spinu, jak i równoczesne ustalenie położenia oraz pędu. Teoria taka spełniłaby marzenie Einsteina o deterministycznym opisie natury bez użycia prawdopodobieństw.

W kontekście tym mówi się często o „ukrytych parametrach", które mają uzupełnić teorię kwantową, aby stała się kompletna. Już w 1932 roku matematyk John (Johann lub János) von Neumann (1903–1957) pokazał jednak w swoim epokowym dziele *Mathematische Grundlagen der Quantenmechanik*, traktującym o matematycznych podstawach mechaniki kwantowej, że ukryte zmienne właściwie nie mogą istnieć. Dowód von Neumanna jest oczywiście poprawny, lecz okazało się, że niezbędne założenia były zbyt wąskie. Zauważyła to już w 1935 roku i odnotowała w swoim dziele *Die naturphilosophischen Grundlagen der Quantenmechanik* (Przyrodniczo-filozoficzne podstawy mechaniki kwantowej) fizyk i filozof Grete Hermann (1901–1984). Niestety jej wyniki pozostały niezauważone. Niezależnie od nich w 1964 roku północnoirlandzki fizyk John Bell (1928–1990) odkrył zbyt ścisłe założenia w dowodzie von Neumanna. Bell nie poprzestał jednak na tym i przy ogólnym założeniu lokalności wyprowadził nierówności, które muszą być spełnione dla wyników oddzielnych pomiarów w przestrzeni, jak w sytuacji EPR. Nierówności te nie odnoszą się do określonych właściwości cząstki (np. jej spinu) czy określonych teorii, lecz jedynie do samych wyników pomiarów, jakby do liczby kliknięć w detektorach. Istotne, że to teoria kwantowa narusza owe nierówności! Z perspektywy tej teorii nie ma w tym nic nadzwyczajnego, ponieważ dotyczy ona przede wszystkim splątanych stanów nielokalnych, czyli nie może spełnić założenia lokalności. Naruszenie to zdumiewa tylko wtedy, gdy wierzy się w rzeczywistość lokalną.

Nierówności Bella można przetestować eksperymentalnie, przy czym eksperymenty te z reguły są przeprowadzane na parach fotonów. Również w tym przypadku mechanika kwantowa świetnie się sprawdziła i lokalna rzeczywistość została obalona. Jeżeli zatem ukryte parametry istnieją, muszą mieć nielokalną naturę. Przykładem jest omówiona poniżej teoria de Broglie'a-Bohma. W ten sposób można obalić też określone nielokalne teorie klasyczne, które przeczą mechanice kwantowej. Stało się to jasne dzięki eksperymentom Zeilingera i grupy jego współpracowników

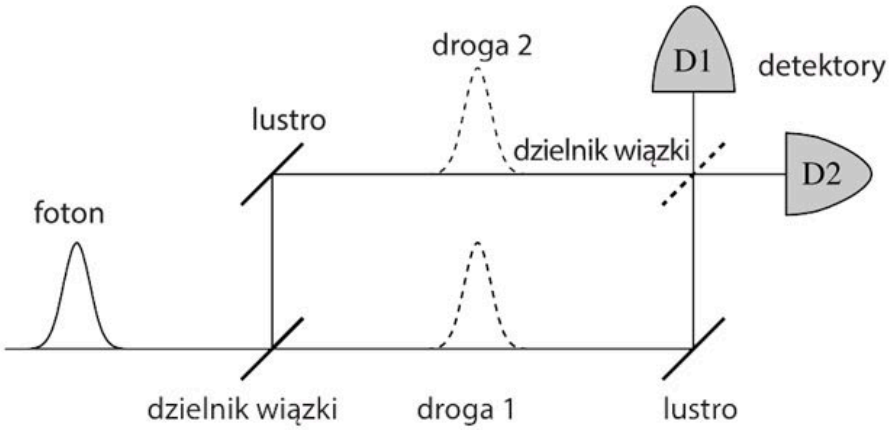
zaprezentowanym w 2007 roku. Nie obaliły one jednak istnienia niezależnej rzeczywistości.

Naruszenie nierówności Bella pokazuje, że układy kwantowe są mocniej skorelowane niż układy klasyczne. Przyczyną tego jest oczywiście splątanie. Za ulubiony przykład klasycznej korelacji służyły Bellowi skarpetki Bertlmanna. Reinhold Bertlmann to wiedeński fizyk znany z tego, że nosi skarpetki różnego koloru. Gdy jedna skarpetka jest czerwona, to druga nie jest czerwona (mamy tu do czynienia z przypadkiem antykorelacji). Dla tej czysto klasycznej sytuacji nierówności Bella oraz założenie lokalnej rzeczywistości, przynajmniej na tym poziomie, są spełnione. Założenie lokalnej rzeczywistości nie jest jednak spełnione, gdy w grę wchodzą splątane stany kwantowe.

Czy można przenieść informację z jednego układu częściowego natychmiast (czyli z prędkością nadświetlną) do drugiego układu częściowego, gdy lokalność zostaje naruszona? Czy mamy tu do czynienia, według określenia Einsteina, z upiornym oddziaływaniem na odległość? W rzeczywistości można udowodnić matematycznie, że nie jest to możliwe, więc teoria względności nie zostanie naruszona. Jest to spowodowane tym, że w sytuacji EPR dwóch obserwatorów, którzy w danym czasie badają pierwszą i drugą cząstkę, mierząc spin na przykład w kierunku pionowym, znajdzie spin w górę lub spin w dół za każdym razem z prawdopodobieństwem ½. Nie mogą zatem rozpoznać, czy drugi obserwator już przeprowadził pomiar spinu i czy spin został ustalony. Po przeprowadzeniu pomiaru teoria kwantowa i tak prognozuje, że każdy kolejny pomiar da ten sam wynik (o ile nie nastąpi zakłócenie). Cała informacja tkwi już w danym globalnym stanie kwantowym, który od początku jest nielokalny i dodatkowe rozchodzenie się informacji z prędkością nadświetlną nie jest możliwe.

Splątanie i nielokalna natura rzeczywistości szczególnie widowiskowo prezentują się w eksperymencie myślowym, który zaproponował John Wheeler w 1981 roku. Eksperyment ten, nazywany eksperymentem „opóźnionego wyboru”, został przeprowadzony w 2006 roku przez

międzynarodową grupę badaczy, nad którą kierownictwo objął Jean-François Roch (Cachan, Francja). Zasada ta została przedstawiona na rysunku 2.3.



2.3. Eksperyment opóźnionego wyboru

Foton trafia na dzielnik wiązki, a następnie biegnie dwiema drogami do dwóch detektorów, odpowiadających tym drogom. Istota eksperymentu polega na tym, aby przed detektorami ustawić, lub nie, drugi dzielnik wiązki, i to dopiero wtedy, gdy foton jest już w drodze (opóźniona decyzja). Żadna informacja, która może przemieszczać się maksymalnie z prędkością światła, nie może dotrzeć z miejsca ustawienia pierwszego dzielnika wiązki do drugiego dzielnika wiązki w chwili przybycia fotonu. Jeśli zdecydujemy się na umieszczenie drugiego dzielnika wiązki, w detektorach pojawiają się zjawiska interferencji (oba detektory reagują z określonym prawdopodobieństwem, podobnie jak w przypadku podwójnej szczeliny – zależnie od kierunku). Foton „przebiegł zatem obie drogi". Jeżeli nie umieścimy drugiego dzielnika wiązki, to jeden albo drugi detektor zareaguje z pięćdziesięcioprocentowym prawdopodobieństwem, a więc nie dojdzie do interferencji prawdopodobieństw. Wygląda to tak, jakby foton biegł tylko jedną albo drugą drogą. Przypuszczalnie foton decyduje się na wybór jednej z dróg albo obu, zanim zapadnie decyzja o umieszczeniu drugiego dzielnika

wiązki. John Wheeler myślał nawet o tym, aby wykorzystać do takiego eksperymentu obie składowe promienia, powstałe w soczewce grawitacyjnej, które przemierzyły kosmiczne odległości.

W rzeczywistości nie ma w tym eksperymencie nic niezwykłego. Staje się on zadziwiający, gdy zostaną w nim zastosowane wyobrażenia fizyki klasycznej. Kwantowa funkcja falowa dla fotonu jest nielokalna i obejmuje obie drogi. Umieszczenie drugiego dzielnika wiązki doprowadzi obie fale składowe do interferencji i nic ponadto. Foton nie jest bowiem klasyczną „kuleczką", która w tajemniczy sposób musi decydować, czy wybiera jedną drogę, czy obie. Wiele z podstawowych eksperymentów w ramach teorii kwantowej przeprowadza się tylko po to, aby udowodnić jej aspekt nielokalności. Sama teoria dotychczas zawsze się sprawdzała.

Informacja kwantowa

Fenomen splątania, tak istotny dla zrozumienia teorii kwantowej, przez wiele lat był uznawany za niezwykły, a nawet paradoksalny element tej teorii. Podejście to zmieniło się dopiero na przestrzeni ostatnich dwudziestu lat. Po pierwsze, dyskusja nad nierównościami Bella i ich eksperymentalne sprawdzenie wzmocniły zaufanie do teorii kwantowej i jej nielokalnej natury. Po drugie, stany splątane bierze się coraz częściej za podstawę w konkretnych zastosowaniach. Chodzi tu o rozwój różnych procesów, które określa się łącznie nieco niefortunnym pojęciem „informacji kwantowej". Stany splątane są wykorzystywane w różnych aspektach informacji kwantowej. Są to zarówno komputery kwantowe, jak i teleportacja kwantowa oraz kryptografia kwantowa.

Klasyczną jednostką informacji jest bit, czyli alternatywa między dwiema możliwościami – między 0 i 1, a zatem prąd płynie albo nie płynie, jest dziura albo jej nie ma. To ostatnie znalazło zastosowanie w płytach CD i demonstruje w widowiskowy sposób, że tak kompleksową strukturę jak IX symfonia Beethovena można zamienić w ciąg bitów.

Jednostką informacji kwantowej jest bit kwantowy, a w skrócie kubit. Bazuje on na układach kwantowych („cząstkach"), które posiadają dwa istotne stany (funkcje falowe) nazywane stanami podstawowymi, tak jak spin w górę lub w dół w przypadku fermionów albo dwa dyskretne poziomy w atomie. Istotne w tej kwestii jest to, że ze względu na zasadę superpozycji można brać pod uwagę dowolne nakładanie się tych stanów podstawowych, czyli w pewnym sensie nakładanie się „0" i „1". Przy dwóch cząstkach istnieją cztery stany podstawowe (odpowiadają kombinacjom 00, 11, 01 i 10), przy trzech cząstkach – osiem i tak dalej. Komputer kwantowy, który pracuje z dowolnymi superpozycjami tych stanów podstawowych, mógłby zatem przy trzech cząstkach zmagazynować osiem zamiast trzech liczb (jak w klasycznym komputerze). Z uwagi na to, że liczba stanów podstawowych rośnie bardzo szybko (wykładniczo) wraz z liczbą cząstek, komputer kwantowy może w zasadzie liczyć o wiele szybciej niż klasyczny komputer, co uzasadnia tak duże zainteresowanie tą kwestią. Ze względu na to, że rozwój w czasie w mechanice kwantowej oddziałuje równocześnie na wszystkie składowe funkcji falowej, komputer kwantowy może liczyć właściwie równocześnie w „wielu światach". Mowa tu o paralelizmie kwantowym.

Doskonałym przykładem, w którym szybkość komputera kwantowego rzeczywiście znajduje zastosowanie, jest tak zwany algorytm Shora. Amerykański informatyk Peter Shor pokazał w 1994 roku, że za pomocą komputera kwantowego możliwa jest faktoryzacja wielkich liczb w krótkim czasie, to znaczy rozłożenie ich na czynniki pierwsze. Wiele dzisiejszych technik szyfrowania (np. w elektronicznych operacjach bankowych lub transakcjach internetowych) opiera się na tym, że mnożenie wielkich liczb jest wprawdzie łatwe, ale proces odwrotny przy obecnych metodach trwa znacznie dłużej. Komputer kwantowy za jednym zamachem podważyłby skuteczność wielu systemów zabezpieczeń. Na razie nie ma bezpośredniego powodu do obaw. Do tej pory udało się skonstruować komputery kwantowe z małą liczbą kubitów (przez magazynowanie naładowanych cząstek), a przeprowadzonych dzięki nim faktoryzacji liczb mogłoby dokonać w pamięci każde dziecko. Główną

przeszkodą w skonstruowaniu takiego komputera jest zaburzający wpływ otoczenia i powodowana przez nie dekoherencja, którą omówimy poniżej jako ogólny fenomen.

Teleportację kwantową i kryptografię kwantową przedstawimy tu jedynie w zarysie. Inaczej niż sugeruje nazwa, podczas teleportacji żaden stan nie zostanie przeniesiony natychmiastowo. Dojdzie raczej do przesłania informacji o stanie kwantowym. W tym celu zmyślnie wykorzystuje się splątanie. Określona informacja musi jednak zostać przekazana „klasycznym kanałem" (jak sieć telefoniczna), co naturalnie nie może się odbyć z prędkością nadświetlną. W kryptografii kwantowej nie chodzi o zaszyfrowanie wiadomości, lecz o bezpieczne przesłanie klucza niezbędnego do jej odczytu. Bezpieczne szyfrowania istnieją już w klasycznej teorii informacji, ale niestety nigdy nie można być pewnym, czy sam klucz nie został już przechwycony. W kryptografii kwantowej powstaje sytuacja podobna do EPR, w której obaj partnerzy (zazwyczaj nazywani Alicją i Bobem) dzielą splątaną parę fotonów. Ewentualny szpieg mógłby zniszczyć splątanie, co Alicja i Bob mogliby odkryć dzięki temu, że w takim przypadku w odróżnieniu od stanu splątania byłyby spełnione nierówności Bella. Wtedy zaprzestano by używać starego klucza i wymyślono nowy.

Za takim rozpoznaniem szpiegostwa kryje się właściwość teorii kwantowej, którą nazywa się twierdzeniem o zakazie klonowania. Można wszak udowodnić, że nie jest możliwe stworzenie dokładnej kopii („klona") stanu kwantowego. Okazuje się bowiem, że klonowanie przeczy zasadzie superpozycji. Twierdzenie nie daje żadnej szansy potencjalnemu szpiegowi. Z powodu tego twierdzenia w komputerze kwantowym nie będzie można również wykonać prostej kopii zapasowej danych. W tym celu trzeba się posłużyć bardziej wymyślnymi metodami. Co ciekawe, okazuje się, że możliwość klonowania pociągałaby za sobą możliwość przesyłania sygnału z prędkością nadświetlną, a więc przez niemożność klonowania uwidacznia się głęboko ukryta harmonia z teorią względności.

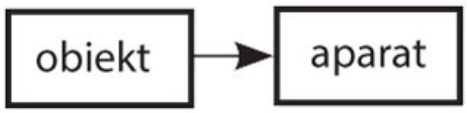
Obecnie można już wytwarzać stany o wysokim stopniu splątania i nie widać końca tego rozwoju. Na przykład w Chinach w 2004 roku wytworzono stan, w którym splątanych było ze sobą aż pięć fotonów. W 2007 roku przy zastosowaniu kryptografii kwantowej z wyspy La Palma na Teneryfę przesłano klucz zabezpieczony przed podsłuchem, a bądź co bądź jest to odległość 144 kilometrów. Zastosowania w informacji kwantowej pokazują, jak ważne (i realne) są splątania. Okazały się one podstawą zrozumienia kwantowego kosmosu.

Jak powstaje klasyczny świat?

Opisany powyżej świat kwantowy początkowo zdaje się nie mieć nic wspólnego z dobrze nam znanym światem fizyki klasycznej. Cząstki nie poruszają się już w czasoprzestrzeni po orbitach, lecz zostają opisane przez funkcje falowe w wielowymiarowej przestrzeni konfiguracyjnej, z których wynika prawdopodobieństwo odnalezienia cząstki podczas „pomiaru" w określonym miejscu. W jaki sposób zatem obserwowany klasyczny świat z jego orbitami wynika z teorii kwantowej? Z uwagi na spójność fizycznego opisu natury powinna istnieć możliwość zrozumienia przypadku granicznego fizyki klasycznej w ramach teorii kwantowej. Relacja fizyki kwantowej do fizyki klasycznej ma także centralne znaczenie dla problematyki interpretacji teorii kwantowej, omówionej w kolejnym rozdziale.

W swojej wspomnianej już książce z 1932 roku von Neumann jako pierwszy opisał oddziaływanie między układem kwantowym a makroskopowym aparatem pomiarowym w modelu matematycznym. Pomiar jest niczym innym jak określonym oddziaływaniem między obiektem a aparatem. W przypadku makroskopowego aparatu zastosowanie znajduje najwyraźniej fizyka klasyczna. Von Neumann nie zakładał jednak tak po prostu zastosowania fizyki klasycznej, lecz roztrząsał akt pomiaru w ramach teorii kwantowej – aparat składa się przecież z atomów, którymi zajmuje się teoria kwantowa. To, co odkrył podczas tego badania, wydawało się całkowicie absurdalne. W opisie

teoretycznym aparat znajdował się na końcu podczas nakładania się na siebie stanów zróżnicowanych pod względem makroskopowym! Jak mogło do tego dojść? Przypatrzmy się najprostszemu przykładowi pomiaru, w którym układ jest jedynie mierzony i nie dochodzi do zakłóceń. Von Neumann określał taką sytuację jako „pomiar idealny" (zob. rys. 2.4).



2.4. Idealny pomiar układu kwantowego aparatem pomiarowym, w którym mogą wystąpić makroskopowe superpozycje

Prostym przykładem jest pomiar spinu w elektronie. Jak już widzieliśmy wcześniej, w odniesieniu do jednego zadanego kierunku (np. kierunek pionowy) możliwe są dwa warianty: spin jest skierowany w górę lub spin jest skierowany w dół. Aby aparat nadawał się do przeprowadzenia pomiaru, musi być tak skonstruowany, że w przypadku pierwszej możliwości (spin w górę) wskazuje inaczej (np. wskazówka wychyla się w lewo) niż w przypadku drugiej możliwości (spin w dół), kiedy wskazówka na przykład wychyla się w prawo. Na razie wszystko wydaje się proste, ale teraz dochodzi jeszcze zasada superpozycji teorii kwantowej. Elektron może przecież znajdować się w stanie, w którym „spin w górę" i „spin w dół" nakładają się na siebie. Zgodnie z regułami teorii kwantowej takie nałożenie się na siebie obu stanów powinno zostać przedstawione na aparacie pomiarowym, a zatem na koniec pomiaru musi on odzwierciedlać superpozycję wskazówki w prawo i wskazówki w lewo, czyli nie będzie pokazywał żadnego definitywnego wyniku pomiaru. Z uwagi na to, że takiego przypadku nigdy nie zaobserwowano i wykraczał on poza wszelkie granice wyobraźni, von Neumann zasugerował nową dynamikę, która unieważnia równanie Schrödingera podczas samego procesu pomiaru. Jest to tak zwany kolaps funkcji falowej. Podczas pomiaru superpozycja (w tym wypadku spinu w górę i spinu w dół oraz

skorelowanej z nimi pozycji wskazówki) miała rozwijać się w sposób indeterministyczny w formie *albo* spinu w górę, *albo* spinu w dół, a więc wskazówka po pomiarze wychyla się albo w lewo, albo w prawo – zależnie od eksperymentu. To, z którym z rezultatów będziemy mieć ostatecznie do czynienia, wynika z prawdopodobieństw wyliczonych na podstawie funkcji falowej.

Problem makroskopowych nałożeń bardzo obrazowo przedstawił Erwin Schrödinger w swoim eksperymencie myślowym z 1935 roku, zastępując aparat żywą istotą – kotem. Od tej chwili kot Schrödingera zdobył pewnego rodzaju popularność w dyskusjach na temat podstaw teorii kwantowej. Na czym polegał ten eksperyment? Najlepiej będzie, jeśli przedstawimy opis tego eksperymentu sporządzony przez samego Schrödingera w 1935 roku:

> Kot zostaje zamknięty w stalowej komorze razem z następującą piekielną maszyną (...): w liczniku Geigera znajduje się maleńka ilość substancji radioaktywnej, tak niewielka, że w ciągu godziny może dojść do rozpadu jednego z atomów, jednak równie prawdopodobne jest też to, że rozpad nie nastąpi. Gdy dojdzie do rozpadu, zadziała licznik i za pośrednictwem przekaźnika uruchomi młoteczek, który rozbije kolbę z cyjanowodorem. Gdy cały układ zostawi się w spokoju na godzinę, można powiedzieć, że kot jeszcze żyje, jeżeli żaden atom się nie rozpadł. (...) Funkcja ψ całego układu wyraziłaby to w taki sposób, że w komorze żywy i martwy kot (...) jest wymieszany lub rozmazany w tych samych proporcjach.

Spin elektronu został w powyższym przykładzie zastąpiony radioaktywnym atomem, a aparat pomiarowy – kotem. Wprowadzony przez von Neumanna kolaps funkcji falowej spowodowałby, że ten niezwykły stan kwantowy „martwy kot – żywy kot" w obserwacji z prawdopodobieństwem wyliczonym z ψ rozpada się na alternatywy „martwy kot" *lub* „żywy kot".

Kiedy więc mamy do czynienia z obserwacją, a kiedy nie? W jednym ze swoich ostatnich wykładów, który odbył się w Princeton 14 kwietnia 1954 roku, Einstein jeszcze raz podjął ten temat. Według Einsteina wydaje się, że świat jest mglistym miejscem, zanim ktoś, na przykład

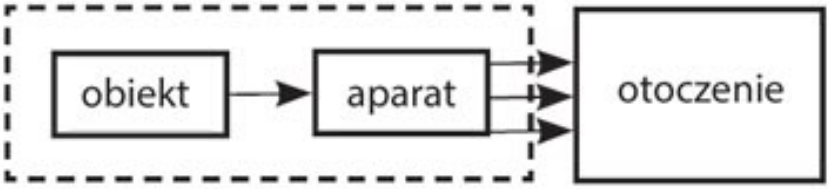
mysz, na niego nie spojrzy. Czy do kolapsu funkcji falowej niezbędna jest zatem żywa istota? A może nawet istota posiadająca świadomość, jak spekulował von Neumann?

Można byłoby po prostu przypuszczać, że świat powstał bez tych zwariowanych superpozycji i istnieją tylko tak zwane „wąskie pakiety falowe", które przypominają ostateczne klasyczne stany. Niestety argument ten nie jest przekonujący, ponieważ nałożenia makroskopowe mogą powstać zawsze z nałożeń mikroskopowych – potwierdza to analiza procesu pomiarowego von Neumanna i przykład z kotem Schrödingera. Przyczyna obserwowalnej klasycznej natury świata musi tkwić gdzie indziej.

W ostatnich dziesięcioleciach nauczyliśmy się, że klasyczny przypadek graniczny można uzasadnić na podstawie teorii kwantowej. Wniosek taki wynika z fundamentalnego badania przeprowadzonego przez H.-Dietera Zeha na Uniwersytecie Ruprechta i Karola w Heidelbergu w 1970 roku. Zeh stwierdził, że w ramach teorii kwantowej pod żadnym pozorem nie można postrzegać obiektów makroskopowych w odizolowaniu, ponieważ są one nierozerwalnie sprzężone z wszechobecnymi stopniami swobody ich otoczenia. Stopnie swobody mogą oznaczać fotony lub molekuły powietrza, które rozpraszają się na tym obiekcie. Bez oddziaływania ze światłem nie moglibyśmy dostrzec żadnych przedmiotów. Obiekty makroskopowe nie mogą zatem spełnić równania Schrödingera, gdyż obowiązuje ono jedynie dla układów izolowanych.

Sprzężenie z otoczeniem zapewne dlatego było tak długo ignorowane, ponieważ w fizyce klasycznej już od czasów Galileusza obserwowano obiekty przede wszystkim w odizolowaniu, wykluczając przy tym otoczenie. W fizyce klasycznej rzeczywiście jest to możliwe i takiemu podejściu zawdzięczamy powstanie mechaniki klasycznej z jej wyidealizowanym pojęciem układu inercjalnego. Jak dowodzą obliczenia ilościowe, w teorii kwantowej nie jest to już możliwe, bo nawet w przypadku niewielkiego sprzężenia z otoczeniem może dojść do splątania stanu układu ze stanem otoczenia bez bezpośredniego zaburzenia

układu. W przeciwieństwie do zaprezentowanego na rysunku 2.4 jedynego oddziaływania układu kwantowego z aparatem pomiarowym bardziej realistyczną sytuację przedstawia raczej rysunek 2.5. Aparat, ponieważ jest makroskopowy, sprzęga się ze stopniami swobody otoczenia. Trzy strzałki między aparatem a otoczeniem mają symbolizować bardzo liczne stopnie swobody otoczenia, czyli wiele fotonów lub molekuł powietrza, które rozpraszają się na aparacie i znikają bez śladu.



2.5. Realistyczne rozszerzenie modelu pomiarowego von Neumanna, w którym uwzględniono otoczenie. Klasyczny proces odbywa się poprzez nieuniknione i nieodwracalne oddziaływanie aparatu z otoczeniem.

Powstaje nielokalny stan kwantowy układu, aparatu i otoczenia. W związku z tym informacja o możliwych makroskopowych nałożeniach aparatu zostaje zdelokalizowana w wyniku uwzględnienia stopni swobody otoczenia. Lokalnie, to znaczy w naszym niewielkim zakątku w wielowymiarowej przestrzeni konfiguracyjnej funkcji falowej, interferencje nie są już wykrywalne. Aparat i tym samym także układ zdają się znajdować w stanie ostatecznym. Prawdopodobieństwo tego stanu można wyliczyć na podstawie mechaniki kwantowej, czyli w rzeczywistości „i” zmieniło się w „lub”. W pewnym sensie aparat pomiarowy staje się klasyczny przez to, że sam jest „mierzony” przez otoczenie. Z uwagi na to, że interferencje kwantowe, nazywane również koherencjami, znikają lokalnie (przy aparacie), proces ten określa się mianem *dekoherencji*. Ma on fundamentalne znaczenie dla zrozumienia teorii kwantowej i jej klasycznego przypadku granicznego. Proces dekoherencji funkcjonuje, ponieważ wszyscy obserwatorzy są lokalni, czyli w przeciwieństwie do stanów splątanych nie zajmują dowolnie odległych

od siebie obszarów. Nielokalny obserwator widziałby całkowity stan kwantowy, zawierający otoczenie, ale nie dostrzegałby już klasycznego świata. Powód istnienia lokalnych obserwatorów pozostaje nadal zagadką.

Tego przejścia do „lub", czyli przejścia do ostatecznego rezultatu pomiaru, w miejscu aparatu (dopóki otoczenie nie jest obserwowane) nie można odróżnić od kolapsu wprowadzonego *ad hoc* przez von Neumanna w 1932 roku. Z tego powodu mówi się o „kolapsie pozornym". Proces dekoherencji, który jest za niego odpowiedzialny, pokazuje, że klasyczne właściwości obiektu nie są jego nieodłącznymi naturalnymi cechami, lecz powstają dopiero wskutek oddziaływania z otoczeniem, czyli są definiowane dopiero przez otoczenie. Oddziaływanie to wyróżnia określone stany aparatu pomiarowego (i innych obiektów makroskopowych) z tego powodu, że koherencje *tych* stanów są tłumione oraz że tłumienie jest wytrzymałe, czyli nie zmienia się w czasie. Takie wyróżnione, klasycznie oddziałujące stany nazywane są „stanami wskaźnikowymi" (co ma budzić skojarzenie ze wskazówką dawnego aparatu pomiarowego).

Być może najważniejszym przykładem dekoherencji jest lokalizacja przedmiotów makroskopowych w przestrzeni. Obiekty takie widujemy wokół nas w określonym położeniu, w dodatku nie niewyraźnie, czyli inaczej niż sugerowałaby teoria kwantowa na podstawie zasady superpozycji. Szczegółowe obliczenia przeprowadzone przez Zeha i Ericha Joosa w 1985 roku pokazują dokładnie, w jaki sposób powstaje taka lokalizacja wskutek oddziaływania ze światłem lub molekułami atmosfery ziemskiej. Nawet przy założeniu istnienia zwariowanej makroskopowej superpozycji (według której na przykład ten sam ołówek leży równocześnie po mojej lewej i prawej stronie) rozproszenie molekuł powietrza postarałoby się o to, aby informacja o tej zwariowanej superpozycji wypłynęła do otoczenia. Lokalnie nie pozostałoby już nic, czyli ołówek leżałby po lewej *albo* po prawej stronie. Dekoherencja taka jest tak efektywna, że wskutek oddziaływania z niezwykle słabym kosmicznym promieniowaniem tła (zob. część o kosmologii) nawet

cząsteczka pyłu w przestrzeni międzygalaktycznej przyjmie właściwości klasyczne. Bezpośrednie zakłócenie cząsteczki pyłu nie jest przy tym konieczne – wystarczy, że promieniowanie tła zna położenie tej cząsteczki i jest splątane z jej stanem, a zatem to raczej otoczenie jest „zakłócone", a nie układ. W pierwszych dziesięcioleciach od powstania teorii kwantowej właśnie ta efektywność nie była doceniana, dlatego dopiero niedawno nauczono się doceniać znaczenie dekoherencji. Rysunek 2.6 przedstawia, w jaki sposób obiekt makroskopowy zostaje zlokalizowany dzięki oddziaływaniu z fotonami.

Oczywiście stopnie swobody otoczenia na ogół również posiadają jakieś otoczenie, ze stanem którego są splątane. Gdzie zatem leży granica? Czy istnieją układy ściśle izolowane? Rzeczywiście łańcuch ten ciągnie się coraz dalej i kończy się w ekstremalnym przypadku na wszechświecie jako całości i na kosmologii kwantowej. Wszechświat jest z definicji jedynym ściśle zamkniętym systemem. Nieuniknione sprzężenie wszystkich systemów ze sobą jest jednym z głównych powodów prowadzenia rozważań na temat kosmosu kwantowego. Więcej o tym w dalszej części książki.

2.6. Rozproszenie fotonów na obiekcie makroskopowym

Czy i w jakim stopniu obiekty zachowują się klasycznie, jest zagadnieniem ilościowym, które można wyliczyć przy danym otoczeniu i danym oddziaływaniu. Z reguły atomy i niewielkie molekuły nie zachowują się klasycznie, ponieważ wpływ otoczenia jest zbyt słaby. Z kolei struktura większych molekuł zostaje zazwyczaj dostatecznie utrwalona przez rozproszenia, więc znany nam z podręczników chemii klasyczny obraz struktur przynajmniej w tym przypadku może się obronić. Z pewnością odnosi się to do tak dużej molekuły, jaką jest substancja genetyczna kwasu dezoksyrybonukleinowego (DNA) – bez jej klasycznej struktury mechanizm dziedziczenia nie mógłby istnieć.

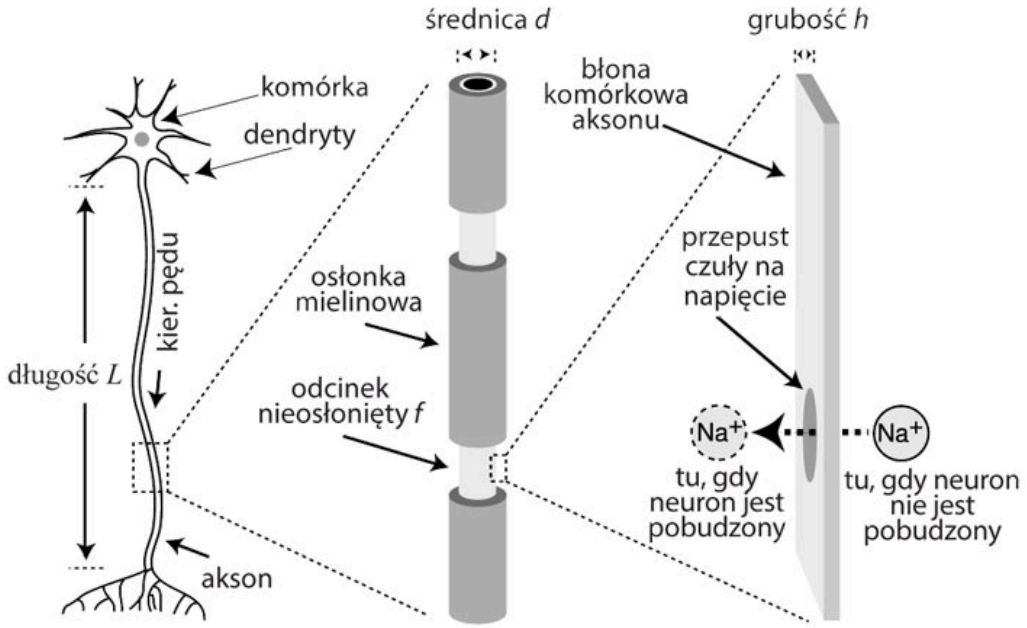
W niektórych sytuacjach wpływ otoczenia może być tak wyraźny, że tłumi każdą czasową zmianę stanu układu. Mówi się wtedy o kwantowym

efekcie Zenona. Zenon z Elei (około 495–około 430 r. p.n.e.), uczeń Parmenidesa, jest znany ze swoich aporii, które miały unaocznić, że w przyrodzie nie dochodzi do zmian. Najbardziej znana aporia traktuje o Achillesie, któremu nie udaje się dogonić żółwia. W teorii kwantowej okoliczności, w których stan zostaje zamrożony, są całkowicie zrozumiałe, a efekt Zenona został eksperymentalnie przetestowany. Efekt ten występuje w układach (atomy, molekuły) z nielicznymi stanami, gdy sprzężenie z otoczeniem jest tak silne, iż permanentna obserwacja tak „unieruchamia" stan, że nie może on ulec zmianie. Stąd porównanie z aporią Zenona.

Jeżeli wpływ otoczenia jest słabszy, znane prawo rozpadu promieniotwórczego, opisujące rozpad jednej substancji w drugą (na przykład istotny w dziedzinie archeologii rozpad węgla-14 w azot-14), może zostać spełnione. Dochodzi do tego tylko wtedy, gdy produkt rozpadu rzeczywiście jest „mierzony" przez otoczenie, czyli nieodwracalnie z nim oddziałuje. W przeciwnym razie produkt rozpadu uległby ponownej interferencji z pierwotnym stanem układu, który nie uległ rozpadowi, i do żadnego rozpadu by nie doszło. Ostatnie można sprawdzić, izolując odpowiednio substancję radioaktywną od otoczenia. W rezultacie udało się potwierdzić wystąpienie interferencji produktu rozpadu z pierwotną substancją!

Wpływ dekoherencji nie jest pożądany w sytuacjach, których celem jest izolowanie układu. Przykładem może być omawiany w ostatnim rozdziale komputer kwantowy. Główną przeszkodą w skonstruowaniu takiego urządzenia jest właśnie wpływ otoczenia, które delokalizuje interferencje niezbędne do poprawnego działania komputera i tym samym niszczy je dla lokalnego obserwatora. Z tego powodu komputery kwantowe składają się jak dotąd zaledwie z niewielu atomów i nie mogą jeszcze wykonywać interesujących obliczeń (jedynie testy rachunkowe, jak mnożenie 3 przez 5). Czy kiedykolwiek uda się wyeliminować dekoherencję, czy też może już zawsze będzie ona uniemożliwiała skonstruowanie efektywnego

komputera kwantowego, to intrygujące pytanie, na które w tej chwili nie znamy jeszcze odpowiedzi.



2.7. Neuron składa się z ciała komórkowego, dendrytów i aksonu. W procesie komunikacji dwóch neuronów impuls elektryczny wędruje wzdłuż aksonu. Różnica między „stanem pobudzonym" a „stanem niepobudzonym" polega przede wszystkim na tym, że po jednej bądź po drugiej stronie błony komórkowej znajduje się około miliona jonów sodu lub potasu.

Niektórzy naukowcy, jak na przykład Roger Penrose, spekulowali, że teoria kwantowa odgrywa znaczącą rolę w zrozumieniu ludzkiej świadomości. W tym kontekście zastosowanie powinny znaleźć przede wszystkim nałożenia stanów różnorodnych pod względem klasycznym. Najczęściej stosowane modele do opisu mózgu mają naturę klasyczną. Opisują one komunikację między neuronami (komórkami nerwowymi) za pośrednictwem tak zwanych sieci neuronowych. W takiej komunikacji impulsy elektryczne wędrują wzdłuż połączeń („aksonów") (porównaj z rys. 2.7). Przywołana przez Penrose'a i innych nieklasyczna superpozycja polegałaby na nakładaniu się na siebie stanów „neuron pobudzony" i „neuron niepobudzony". Dla niezależnego obserwatora superpozycja taka

nie miałaby żadnego sensu. Z uwagi na to, że neurony w mózgu nie są odizolowane (oddziałują z molekułami wody i jonami), należy oczekiwać, że ich otoczenie natychmiast doprowadzi do dekoherencji: neuron pobudzony lub neuron niepobudzony. W 2000 roku szwedzko-amerykański fizyk Max Tegmark opublikował na ten temat szczegółowe obliczenia. Odkrył, że dekoherencja działa na skali czasu, która jest o wiele rzędów wielkości mniejsza niż skala czasu pobudzenia neuronów i procesów pokrewnych. Ostatnia jest stosowana w zakresie milisekund, a pierwsza – w zakresie 10^{-20} sekundy. Dekoherencja przebiega więc nieporównywalnie szybciej. Klasyczne modele są zatem odpowiednie do opisu funkcjonowania mózgu, chociaż ich uzasadnienie wynika dopiero z teorii kwantowej.

Od 1996 roku istnieje wiele eksperymentów, które z powodzeniem wyjaśniają przebieg w czasie procesu dekoherencji i tym samym sprawdzają prognozy teorii kwantowej także w tym aspekcie. Na przykład wspomniana już w rozdziale 2 grupa wiedeńska z Antonem Zeilingerem na czele wykorzystała swoje eksperymenty interferencji z fulerenami również w celu obserwacji stopniowej utraty możliwości interferencyjnych przez dekoherencję. Co pełni funkcję otoczenia w tych eksperymentach? Może to być rozproszenie z innymi molekułami bądź emitowanie fotonów przez fulereny podczas podgrzewania. Istotne jest tylko to, aby doszło do splątania z oddziałującymi stopniami swobody, które „uciekają”, a zatem nie istnieją już lokalnie (w przypadku fulerenów). Wyemitowane podczas podgrzewania fotony znajdującej się w superpozycji molekuły w nieodwracalny sposób wynoszą informację o tej superpozycji. W tych wyszukanych eksperymentach można obserwować dekoherencję w czasie. W makroskopowym świecie natomiast zmiany zachodzą tak szybko, że powstaje wrażenie „skoku kwantowego”. Pojęcie to, które zadomowiło się już w reklamach, w rzeczywistości wprowadza w błąd (jak niejedna reklama), bo gdy dokładnie przyjrzymy się temu zjawisku, nie zauważymy żadnego skoku.

Z samej uniwersalności teorii kwantowej wynika, kiedy i w jaki sposób powstają właściwości klasyczne oraz dlaczego codzienny świat *sprawia wrażenie* klasycznego. To proces dekoherencji uzasadnia pojawienie się świata klasycznego. W normalnych okolicznościach odpowiednie czasy trwania dekoherencji są tak krótkie, że moment przejścia jest niezauważalny – jest to możliwe jedynie dzięki zastosowaniu skomplikowanych eksperymentów. Relacja teorii kwantowej do jej klasycznego przypadku granicznego odgrywa ważną rolę w kwestii odpowiedzi na pytanie o poprawną interpretację teorii kwantowej – więcej na ten temat poniżej.

Interpretacje teorii kwantowej

Teoria kwantowa jest bardzo skuteczną teorią i jak dotąd nie było eksperymentu, który rzuciłby choćby cień wątpliwości w kwestii jej obowiązywania. Wszyscy fizycy są zgodni co do zastosowania formalizmu teorii kwantowej, mimo to nadal toczą się burzliwe dyskusje wokół znaczenia tej teorii i jej interpretacji. Jaka jest tego przyczyna?

W ostatnim rozdziale przedstawiliśmy problem makroskopowych superpozycji i przykład eksperymentu z kotem Schrödingera, który w myśl teorii kwantowej może być jednocześnie żywy i martwy. Interpretacje teorii kwantowej odnoszą się przede wszystkim do tego właśnie problemu.

Za podręcznikową interpretację uznaje się niekiedy tak zwaną kopenhaską interpretację teorii kwantowej. Powstała ona w wyniku intensywnych dyskusji prowadzonych między Bohrem, Heisenbergiem a innymi naukowcami w Kopenhadze w latach 1925–1927. Nie istnieje jednak żadna dokładna definicja tej interpretacji, być może dlatego, że nawet Bohr i Heisenberg trochę rozmijali się w swoich poglądach. Z tego powodu interpretacja ta sprawia niekiedy wrażenie nie całkiem spójnego zlepku różnych wersji. Centralnym pojęciem jest w każdym razie komplementarność. Pojęcie to po raz pierwszy zaproponował Bohr na konferencji we włoskiej miejscowości Como we wrześniu 1927 roku. Komplementarność początkowo odnosiła się tylko do zachowań obiektów

kwantowych, które wykazywały cechy cząstek lub fal, lecz nigdy obu równocześnie. Fale i cząstki są komplementarne. Bohr uczynił później z komplementarności główne pojęcie w swojej filozofii, która daleko wykracza poza samą fizykę. Heisenberg nie mógł się przekonać do komplementarności, ponieważ nie chciał w teorii fizycznej sprzecznych pojęć, takich jak obraz fal czy obraz cząstek. Język matematyki jest zawsze jednoznaczny. Heisenberg pisał w 1930 roku w swojej książce na temat teorii kwantowej, że w przypadku chęci przejścia od matematyki do obrazowego opisu procesów trzeba zadowolić się niepełnymi analogiami, jakie dają nam obraz fal czy obraz cząstek. Jak już się przekonaliśmy w poprzednim rozdziale, na podstawie formalizmu teorii kwantowej można zrozumieć, kiedy ten sam obiekt zachowuje się jak cząstka, a kiedy jak fala. Pojęcie komplementarności jest zatem zbędne na podstawowej płaszczyźnie. Stan można jednoznacznie opisać przez funkcję falową w wielowymiarowej przestrzeni konfiguracyjnej, która w normalnej trójwymiarowej przestrzeni zależnie od sytuacji zachowuje się jak cząstka albo normalna fala.

Oprócz komplementarności kopenhaską interpretację charakteryzują jeszcze dwie właściwości, a mianowicie: konieczność stosowania w opisie procesu pomiarowego pojęć fizyki klasycznej oraz wprowadzenie formalnej redukcji funkcji falowej bez znaczenia dynamicznego. Fizyka klasyczna, jak twierdził Bohr, jest *a priori* niezbędna, aby opisać aparaty pomiarowe, i to pomimo tego, że aparaty składają się z atomów. Jako jeden z twórców interpretacji kopenhaskiej w dyskusjach z Einsteinem, nie cofnął się on nawet przed zastosowaniem zasady nieoznaczoności w odniesieniu do aparatów pomiarowych, aby uchronić je przed atakami ze strony Einsteina.

W rzeczywistości jednak wraz z upływem czasu Bohr zmienił istotne punkty w swojej interpretacji. Przed wynikami prac EPR z 1935 roku zawsze łączył on z pomiarem niekontrolowane zakłócenie podlegającego pomiarowi układu wskutek oddziaływania aparatu pomiarowego. Po publikacji pracy EPR nie było to już możliwe. EPR przenieśli przecież

wnioski wynikające z pomiaru jednego układu na układ (druga cząstka), który jest bardzo oddalony i zgodnie z założeniem nie podlega zakłóceniom. W swojej odpowiedzi na prace EPR Bohr zajął mocno operacjonistyczne stanowisko, które przysłoniło kluczowe pojęcie splątania. Odrzucał on po prostu fakt istnienia układów kwantowych, dopóki nie oddziałują one z aparatem pomiarowym. W ten sposób uniknął sideł zastawionych przez EPR. Czy jednak odrzucenie rzeczywistości mikroskopowej nie jest zbyt wysoką ceną za utrzymanie interpretacji?

W interpretacji kopenhaskiej rola pojęcia redukcji czy kolapsu funkcji falowej nie jest do końca jednoznaczna. W 1932 roku von Neumann postulował kolaps jako proces dynamiczny, który odpowiada konkretnej zmianie funkcji falowej w rzeczywistości. W odróżnieniu od niego większość zwolenników interpretacji kopenhaskiej traktowała kolaps jako czyste pomnażanie wiedzy i przeczyła tym samym istnieniu funkcji falowej jako części rzeczywistości. Według nich miała ona wyłącznie opisywać przejście od potencjalności świata kwantowego do aktualności świata klasycznego, bez własnego udziału w dynamice. Wyobrażenie rzeczywistego kolapsu byłoby trudno pogodzić z ideą komplementarności Bohra, ponieważ akurat fizyczna analiza procesu pomiarowego jest w tym przypadku wykluczona. Niezaprzeczalną wadą interpretacji kopenhaskiej jest właśnie fakt, że ukrywa się ona za niejasnymi pojęciami i nie dopuszcza do siebie żadnych pytań. Nic więc dziwnego, że powstały alternatywne interpretacje.

W poprzednim rozdziale przedstawiliśmy szczegółowo, w jaki sposób nieodwracalny proces dekoherencji prowadzi do tego, że obserwator nie może lokalnie zauważyć zadziwiających superpozycji, jak na przykład tej z martwym i żywym kotem Schrödingera. Proces ten można opisać dzięki formalizmowi teorii kwantowej i jest on niezależny od jej interpretacji, choć znaczenie, jakie przypisuje się dekoherencji, zależy od preferowanej interpretacji.

Opis dekoherencji odbywa się w ramach teorii kwantowej i dlatego nie narusza jej liniowości. Wprawdzie wyjaśnia on, dlaczego nie można

lokalnie zaobserwować nałożeń makroskopowych, ale są one mimo to obecne w całym układzie. Na przykładzie kota Schrödingera oznacza to, że całościowy stan zawiera nałożenia „martwy kot – obserwator widzi martwego kota – otoczenie jest skorelowane z martwym kotem" oraz „żywy kot – obserwator widzi żywego kota – otoczenie jest skorelowane z żywym kotem". Aby otrzymać z tego tylko pierwszą *lub* drugą składową, trzeba najprawdopodobniej uzupełnić formalizm teorii kwantowej o dynamiczny kolaps (redukcję), jednak w rzeczywistości nie ma takiej konieczności. W 1957 roku Hugh Everett (1930–1982) w swojej pracy doktorskiej, której promotorem był John Wheeler, postulował po prostu, że taki kolaps nie istnieje, równanie Schrödingera jest uniwersalne, a wszystkie składowe mikroskopowej i makroskopowej superpozycji istnieją w rzeczywistości. Występuje zatem wersja obserwatora, który widzi martwego kota, i wersja obserwatora, który widzi żywego kota. Być może wpływ na Everetta wywarł, przynajmniej częściowo, wspomniany już wykład Einsteina z 14 kwietnia 1954 roku, w którym omówił on problem nałożeń makroskopowych i wynikającej z nich, jak twierdził Einstein, niekompletności teorii kwantowej. W swej pracy doktorskiej Everett przedstawił interpretację, która zakładała kompletność teorii kwantowej. Formalizm miał dostarczyć własną interpretację, a teoria decydować o tym, co jest obserwowalne. Zgodnie z interpretacją Everetta wszystkie możliwe wyniki pomiarów istnieją równocześnie, łącznie z przypisanymi im kopiami obserwatorów. Z tego powodu interpretację tę nazwano później interpretacją wielu światów, co jednak jest nieco mylące, ponieważ w rzeczywistości istnieje tylko jeden świat kwantowy. Odgałęzienia, występujące podczas pomiaru, odpowiadają tylko drodze obserwatora, którego lokalność jest zadana przez funkcję falową. Jednocześnie znika propagowany przez Johna von Neumanna paralelizm psychofizyczny, będący bezpośrednim odpowiednikiem stanu układu i stanu postrzegania obserwatora, gdyż w nielokalnym świecie kwantowym istnieje wiele wersji tego samego lokalnego obserwatora. To właśnie ten aspekt interpretacji napotyka na gwałtowny opór, bo jak można twierdzić,

że istnieje wiele wersji obserwatora, skoro nigdy nie dostrzeżono więcej niż jednej z nich?

Everettowi brakowało jeszcze jednego ważnego elementu do stworzenia spójnej interpretacji, a mianowicie dekoherencji. Dopiero ona pozwala zrozumieć, dlaczego nie można dostrzec innych makroskopowych składowych funkcji falowej. Powodem jest to, że informacja o nich uciekła, wchodząc w korelacje ze stopniami swobody otoczenia. Interpretacja Everetta łącznie z procesem dekoherencji dostarcza spójnego i kompletnego sposobu postrzegania teorii kwantowej. Tym samym zostaje rozwiązany problem „granicy Heisenberga", dotyczący miejsca wyznaczenia granicy między opisem za pomocą funkcji falowej a klasycznym opisem obrazowym. Granica ta odpowiada właśnie temu miejscu, w którym w wyniku oddziaływania z otoczeniem pojawia się dekoherencja. Jest to kwestia, którą w konkretnej sytuacji można wyjaśnić, korzystając z badań ilościowych. Po przyjęciu pewnych dodatkowych założeń (które jednak nadal są przedmiotem dyskusji) można również przygotować interpretację prawdopodobieństwa teorii kwantowej. W tym wypadku prawdopodobieństwa mają sens tylko wtedy, gdy są interpretowane jako względna częstość wyników pomiarów. Interpretacja prawdopodobieństwa na podstawie interpretacji Everetta wynika z założenia, że nie żyjemy w „nietypowej" części pełnego nielokalnego świata kwantowego. Może właśnie z tego powodu Carl Friedrich von Weizsäcker (1912–2007) określił interpretację Everetta jako jedyną, która nie sięga wstecz w stosunku do tego, co udało się już zrozumieć dzięki teorii kwantowej, i wykracza poza te ramy.

Czy w kontekście światopoglądowym nie płacimy jednak zbyt wysokiej ceny za taką interpretację? Niekoniecznie. Przypomnijmy sobie, co Goethe pisał w swoich *Refleksjach i maksymach* na inny temat:

> To, co nas tak bardzo zbija z tropu, gdy chcemy ideę poprzeć zjawiskiem, to fakt, że często i zazwyczaj zaprzecza ona zmysłom. System kopernikański opiera się na idei, którą trudno było pojąć i która nadal codziennie zaprzecza naszym zmysłom...

Co ciekawe, także Everett wskazuje w swojej pracy na analogię do przewrotu kopernikańskiego. Zmysłom ukazuje się Słońce okrążające Ziemię, podczas gdy dopiero fundamentalna teoria (mechanika Newtona) ujawnia prawdziwy stan rzeczy. W historii nauki wiara w podstawowe równania i prostotę teorii prawie zawsze sprawdzała się lepiej niż trzymanie się prostoty faktów. Dlaczego zatem mielibyśmy od tego odchodzić w przypadku teorii kwantowej? Być może właściwy przewrót kopernikański jest jeszcze przed nami?

Prostota świata kwantowego ukazuje się tylko niejako „z zewnątrz", z perspektywy lotu ptaka fizyka teoretycznego, który formułuje równania na papierze (porównanie fizyka Maxa Tegmarka). Natomiast z perspektywy żaby rzeczywistego obserwatora w świecie kwantowym świat ten sprawia wrażenie skomplikowanego. Deterministycznie rozwijający się świat kwantowy ukazuje się, patrząc z perspektywy naszego małego zakątka, jako indeterministyczny i opanowany przez prawa rachunku prawdopodobieństwa.

Wielokrotne i równoczesne istnienie tego samego makroskopowego obserwatora wydaje się jednak tak nieprawdopodobne, że nie dziwi poszukiwanie alternatyw. Dość osobliwa interpretacja, choć nie tak odległa od punktu widzenia Everetta, opiera się na pracach Louisa de Broglie'a i Davida Bohma (1917–1992). W skróconej formie funkcjonuje ona pod nazwą „interpretacja Bohma". W rzeczywistości chodzi raczej o odrębną teorię, a nie o interpretację. Obok funkcji falowej wprowadza się w niej dodatkowo zmienne, które opisują klasyczne cząstki. W związku z roztrząsanymi powyżej nierównościami Bella można byłoby pokusić się o nazwanie ich „ukrytymi zmiennymi". Jak celnie zauważył sam Bell, chodzi jednak bardziej o wyróżnione zmienne, gdyż są to „cząstki", które są obserwowane. Teoria Bohma jest w stanie odtworzyć prognozy mechaniki kwantowej. Oczywiście rzeczywistość opisana teoretycznie musi być przy tym nielokalna, w przeciwnym razie zostałyby spełnione nierówności Bella, a to stałoby w sprzeczności z eksperymentem. Można to osiągnąć, gdy funkcja falowa staje się

rodzajem pola prowadzącego („fali pilotującej”) dla cząstek. Ich prędkości nie są już (jak u Newtona) podawane dowolnie, lecz ustalane przez funkcję falową, która nadal spełnia założenia równania Schrödingera. Ważne jest, aby początkowy rozkład cząstek był określany przez kwadrat funkcji falowej, ponieważ tylko wtedy możliwa jest zwykła interpretacja prawdopodobieństwa teorii kwantowej. Tuż po opublikowaniu swoich prac z 1952 roku na ten temat Bohm starał się o wsparcie znanych fizyków, ale bezskutecznie. Na przykład Pauli w liście z 6 stycznia 1952 roku do Markusa Fierza (1912–2006) stwierdził, że teoria Bohma jest klasycznym deterministycznym mitem atomowych zdarzeń, w których rzeczywistość rozpada się na dwie znajdujące się obok siebie odrębne połowy: cząstkę i falę pilotującą. (Wspominał także o dowcipach, które opowiadano sobie w 1927 roku: o „niani” [fali] prowadzącej „dziecko” [cząstkę] za rękę).

Jak teoria Bohma odnosi się do zagadnienia makroskopowych nałożeń, czyli problemu kota Schrödingera? W sytuacjach odpowiadających „pomiarowi” (czyli określonej klasie oddziaływań) dochodzi nadal do nałożeń odpowiednich składowych funkcji falowej. „Cząstka” (może nią być również kot Schrödingera) zostaje jednak uwięziona w jednej ze składowych, podczas gdy inna składowa jest pustym pakietem falowym, który zwykle oddala się od składowej niosącej cząstkę i nie powraca. Pod tym względem teoria Bohma przypomina interpretację Everetta lub, jak wyraził to Bell, Everett jest jak Bohm bez orbit cząstek.

Teoria Bohma jak dotąd napotyka przeszkody, gdy chodzi o opis układów z nieskończenie wieloma stopniami swobody – układy takie są istotne w fizyce cząstek elementarnych (zob. następny rozdział). Niezależnie od tego, czy w przyszłości uda się rozwiązać ten problem, teoria Bohma zdaje się dźwigać zbędny balast w postaci dodatkowych cząstek. To, kiedy i w jaki sposób układy wykazują cechy cząstek (są lokalizowane), wynika tak czy inaczej ze zwykłego sformułowania teorii kwantowej w odniesieniu do procesu dekoherencji. Po co zatem postulować istnienie dodatkowych cząstek?

Prawdziwe alternatywy dla teorii kwantowej postulują zmianę równania Schrödingera. Nowe równanie miałoby być tak skonstruowane, aby w eksperymentalnie sprawdzonych sytuacjach zgadzało się z równaniem Schrödingera, ale prowadziło do ostatecznego wyniku pomiarowego w sytuacjach makroskopowych. Z początkowego nakładania się na siebie wielu składowych w funkcji falowej pozostaje więc w sposób indeterministyczny jedna składowa. Kot Schrödingera jest wtedy żywy *lub* martwy i w przeciwieństwie do interpretacji Everetta na całym świecie istnieje tylko jedna składowa z martwym lub żywym kotem. Są to teorie, które rzeczywiście opisują kolaps (redukcję) funkcji falowej w układzie dynamicznym, nie tylko słownie (jak w interpretacji kopenhaskiej), lecz także za pomocą równań. Powstaje zatem kwestia natury empirycznej, dotycząca tego, czy to alternatywne teorie kolapsu są fundamentalnie poprawne, czy też równanie Schrödingera. Dotychczas równanie Schrödingera zawsze się sprawdzało. Jest to związane głównie z tym, że proces dekoherencji symuluje już kolaps pozorny i dlatego prawdziwy kolaps trudno jest zidentyfikować. Teorie kolapsu są jednak jedyną prawdziwą alternatywą dla interpretacji Everetta.

Ze względu na obecną eksperymentalną nierozróżnialność wybór między jedną z alternatyw, czyli zwykłą teorią kwantową według interpretacji Everetta, a jedną z teorii kolapsu uznaje się za kwestię gustu. W praktycznym zastosowaniu nie widać żadnej różnicy. Od wybranej alternatywy zależy jednak w znacznym stopniu dalszy kierunek rozwoju fizyki. Dotyczy to w szczególności opisanej w kolejnej części kosmologii kwantowej. W każdym razie bezpieczniej jest trzymać się empirycznie potwierdzonej teorii kwantowej wraz z równaniem Schrödingera, przynajmniej dopóki wyniki eksperymentów nie zmuszą nas do zmiany zdania.

Teoria kwantowa i cząstki elementarne

Teoria kwantowa została pierwotnie sformułowana jako teoria atomów i ich oddziaływań ze światłem. Z czasem zrozumiano, że jej skuteczność

sięga o wiele dalej. Na przykład precyzyjnie opisuje ona zachowanie jąder atomów i ich części składowych. Jądra atomów składają się z dodatnio naładowanych protonów i elektrycznie neutralnych neutronów, które łącznie określa się mianem nukleonów (cząstki jądrowe). Są one około 2000 razy cięższe od elektronu. Same nukleony składają się z kolei z innych cząstek, nazywanych kwarkami i gluonami. Gluony pełnią funkcję cząstek wiążących, odpowiedzialnych za przyciąganie kwarków między sobą, co utrzymuje je w jądrach. Kwarki, oprócz zwykłego ładunku elektrycznego, niosą nowy rodzaj ładunku, który metaforycznie określa się mianem ładunku kolorowego, choć z kolorami nie ma nic wspólnego. Gluony sprzęgają się z tym ładunkiem kolorowym, analogicznie do fotonów sprzęgających się z ładunkiem elektrycznym. Nowością jest jednak to, że gluony same niosą ładunek kolorowy w przeciwieństwie do elektrycznie neutralnego fotonu. Gluony oddziałują zatem także między sobą, nie tylko z kwarkami. Gluon, podobnie jak foton, posiada spin = 1, a kwarki analogicznie do protonu – spin = ½, czyli są fermionami. Zasięg oddziaływania kolorowego jest znacznie słabszy niż oddziaływania elektromagnetycznego, lecz przewyższa go swoją mocą, dlatego zwany jest *oddziaływaniem silnym*. Zarówno nukleony, jak i ich części składowe, czyli kwarki i gluony, można opisać w ramach teorii kwantowej za pomocą funkcji falowych i podobnie jak w fizyce atomowej zastosowanie znajduje interpretacja prawdopodobieństwa. Powszechnie stosowanym pojęciem są „cząstki elementarne", niezależnie od tego, czy cząstki te składają się z innych cząstek (jak nukleony), czy też nie (jak zgodnie z obecnym stanem wiedzy kwarki i gluony).

Oprócz opisanych już oddziaływań (grawitacja, elektromagnetyzm, oddziaływanie silne) istnieje jeszcze jedno oddziaływanie potwierdzone empirycznie – tak zwane *oddziaływanie słabe*, słabsze niż elektromagnetyzm i oddziaływanie silne, ale istotne w przypadku wielu fenomenów. Leży ono u podstaw radioaktywności i ma zasadnicze znaczenie w niezbędnej dla naszego życia reakcji syntezy jądrowej na Słońcu.

Omawiana w poprzednich rozdziałach mechanika kwantowa zajmuje się zazwyczaj sytuacjami, w których występujące prędkości są o wiele mniejsze niż prędkość światła. Z tego powodu nie ma konieczności stosowania szczególnej teorii względności. W celu otrzymania bardziej fundamentalnego opisu należy jednak odpowiednio uogólnić teorię kwantową, ponieważ w fizyce cząstek elementarnych wysokie prędkości nie należą do rzadkości. Brytyjski fizyk Paul Dirac (1902–1984) sformułował w 1928 roku równanie dla cząstek elementarnych o spinie = ½ (wtedy myślano tylko o elektronie), które w taki sposób uogólnia równanie Schrödingera, że szczególna teoria względności zostaje uwzględniona automatycznie. „Równanie Diraca" miało zdumiewające konsekwencje, gdyż przewidywało, że obok elektronu istnieje jeszcze inna cząstka o tej samej masie, ale o przeciwstawnym (czyli dodatnim) ładunku, którą ochrzczono mianem pozytonu. Ogromnym sukcesem teorii Diraca było odkrycie w 1932 roku pozytonu na płytach fotograficznych wystawionych na działanie promieniowania kosmicznego (wysokoenergetyczne cząstki, które docierają do nas z kosmosu). Pozytonem nazywa się przynależną do elektronu antycząstkę. Wszystkie cząstki mają w rzeczywistości przeciwstawnie naładowane antycząstki, przy czym w pojedynczych przypadkach cząstek neutralnych antycząstki mogą być z nimi identyczne. Antyproton i antyneutron zostały odkryte w 1955 roku w kalifornijskim Berkeley. Antyneutron różni się od neutronu mimo jego neutralności, ponieważ neutron składa się z trzech elektrycznie naładowanych kwarków, których ładunek sumuje się wprawdzie do zera, lecz przy odbiciu ładunkowym (zmianie znaków ładunków) naprowadza na trzy inne kwarki.

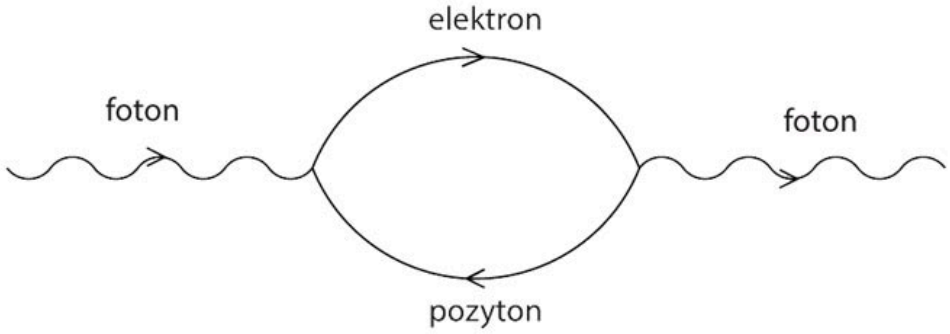
Uwzględnienie szczególnej teorii względności w teorii kwantowej ma jednak dużo poważniejsze konsekwencje. Okazuje się bowiem, że wszystkie „cząstki" należy opisać na poziomie fundamentalnym za pomocą teorii pola. Pole jest wielkością, która przypisuje każdemu punktowi w czasoprzestrzeni określoną wartość. Z fizyki klasycznej znane są pojęcia pola elektromagnetycznego i pola grawitacyjnego. Z tego względu, że w czasoprzestrzeni istnieje nieskończenie wiele punktów,

teoria pola posiada nieskończenie wiele stopni swobody. Prawdziwa cząstka ma natomiast tylko cztery stopnie swobody – trzy kierunki w przestrzeni oraz czas. Jak się okazuje, oddziaływań między prawdziwymi cząstkami (w odróżnieniu od pól) nie można formułować w taki sposób, aby podczas przejścia z jednego układu inercjalnego do innego pozostały one niezmiennicze. Niezmienniczość taka uderzająco zaprzeczałaby teorii względności.

Teoria kwantowa jest zatem na fundamentalnej płaszczyźnie zawsze kwantową teorią pola, czyli teorią kwantową z nieskończenie wieloma stopniami swobody. Z tego powodu jest także bardziej ogólna niż równanie Diraca. Naturalnie istnieje wiele sytuacji, w których można pominąć aspekt pola i prowadzić obliczenia, „opierając się na mechanice kwantowej". W przeciwnym razie równania Schrödingera i Diraca nie byłyby tak skuteczne. Kwantowa teoria pola przewiduje ponadto nowe zjawiska. Na przykład elektron i pozyton mogą się wzajemnie zniszczyć i przekształcić w dwa fotony, albo odwrotnie – dwa fotony mogą się przekształcić w parę elektron-pozyton. Proces taki nazywa się kreacją par, a proces odwrotny w czasie – anihilacją par. Kreacja par jest możliwa zawsze wtedy, gdy do dyspozycji zostanie oddana energia w formie energii promieniowania, odpowiadająca co najmniej masie spoczynkowej pary. Na podstawie teorii kwantowej można wyliczyć wtedy określone prawdopodobieństwo dla kreacji pary. W polu elektrycznym także istnieje pewne prawdopodobieństwo kreacji pary elektron-pozyton, przy czym niezbędna do tego energia znika później z pola elektrycznego.

Kwantowa teoria pola zakłada możliwość, że cząstki istnieją tylko przez krótki moment, jakby wirtualnie. Wskutek zasady nieoznaczoności Heisenberga pomiędzy energią a czasem cząstki mogą „pożyczyć" sobie na pewien czas energię, która jest im potrzebna do zaistnienia. Foton może zamienić się w wirtualną parę elektron-pozyton i natychmiast wrócić do swojej pierwotnej formy. Występująca przez krótki czas nieoznaczoność energii wystarcza, aby uwolnić energię tkwiącą w masie spoczynkowej pary. Amerykański fizyk Richard Feynman (1918–1988) stworzył dla

takich procesów łatwy do zapamiętania język znaków, określany mianem „diagramów Feynmana”. Rysunek 2.8 pokazuje diagram Feynmana dla wirtualnego istnienia pary elektron-pozyton. Za takimi diagramami kryją się w rzeczywistości wzory z zakresu teorii kwantowej, które umożliwiają dokładne obliczenia zarówno prawdopodobieństwa, jak i właściwości interferencyjnych owych procesów. Z uwagi na to, że swobodnie poruszający się w próżni foton może stale przekształcać się w takie pary elektron-pozyton, czyli może w pewnym sensie „polaryzować” próżnię, nazywamy ów proces także polaryzacją próżni. Z kolei swobodnie poruszający się elektron może wyemitować wirtualny foton i wchłonąć go ponownie po krótkiej chwili. Wszystkie te procesy tak zwanej elektrodynamiki kwantowej mają wpływ na widma atomowe, zmieniając w niewielkim stopniu wartości energii. Są to zjawiska, które można zaobserwować eksperymentalnie.



2.8. Przykład diagramu Feynmana

Dotychczas obserwowane cząstki i ich oddziaływania są ujęte w modelu standardowym fizyki cząstek elementarnych. Za cząstki podstawowe (czyli takie, które zgodnie z dzisiejszą wiedzą nie są zbudowane z innych cząstek) uznaje się kwarki, leptony oraz cząstki związane z tak zwanymi polami cechowania. O kwarkach była już mowa. Leptony są cząstkami o spinie = ½, które nie uczestniczą w oddziaływaniu silnym. Ich nazwa pochodzi od greckiego słowa oznaczającego „lekki”. Najważniejszymi przedstawicielami leptonów są elektron i jego antycząstka, pozyton.

W grupie tej znajdują się jeszcze dwie cząstki (wraz z antycząstkami), które różnią się od elektronu jedynie masą. Cząstki mion i taon są cięższymi wersjami elektronu. Mion jest 400 razy cięższy, a taon – 3500 razy. Taon jest prawie dwukrotnie cięższy od protonu, więc z pewnością nie jest już „lekką" cząstką (określenie „lepton" zostało jednak wprowadzone przed odkryciem taonu w 1975 roku).

W grupie leptonów znajduje się jeszcze dostojna gromada cząstek nazywanych neutrinami. To cząstki nienaładowane, które biorą udział tylko w oddziaływaniach słabych (z wyjątkiem grawitacji). Przez długi czas uważano, że nie mają masy. Dopiero w latach dziewięćdziesiątych XX wieku, podczas obserwacji kwantowego procesu interferencji, okazało się, że muszą one jednak posiadać niewielką masę, której wartość nie jest jeszcze znana. Występują trzy rodzaje neutrin (wraz z antycząstkami), a każdy z nich w naturalny sposób przyłącza się odpowiednio do elektronu, mionu i taonu. Ze względu na to, że kwarki również istnieją w trzech podstawowych wersjach (wraz z antycząstkami), które w pewnym sensie pasują do trzech „elektronów" i trzech neutrin, mówi się o trzech rodzinach w modelu standardowym. Jak dotąd brakuje wyjaśnienia, skąd wzięła się liczba trzy. Otwarte pozostaje przede wszystkim pytanie, czy istnieje związek z trzema wymiarami przestrzennymi naszego świata.

Neutrina przeżyły interesujące narodziny. Wolfgang Pauli wywnioskował ich istnienie w 1930 roku na podstawie przesłanek czysto teoretycznych. Gdyby jednak nie istniały, to w przypadku określonego procesu (rozpadu neutronu na proton i elektron) oznaczałoby to naruszenie prawa zachowania energii – cóż by to była za okropność. Z tego powodu Pauli postulował, że podczas rozpadu neutronu musi dodatkowo powstawać niewidoczna wtedy cząstka, która przekazuje brakującą energię, czyli neutrino elektronowe (a dokładniej jego antycząstka). Zostało ono odkryte dopiero w 1956 roku. Kolejne rodzaje neutrin, czyli neutrino mionowe i taonowe, odkryto kolejno w 1962 i 2000 roku.

Za czasów Pauliego prognozy co do istnienia nowej cząstki wymagały jeszcze pewnej odwagi. Dziś na papierze widnieje wiele hipotetycznych

cząstek, dla których brak jeszcze jakiegokolwiek eksperymentalnego punktu zaczepienia. Szczególnie znane są cząstki tak zwanej supersymetrii (w skrócie: SUSY). Postuluje się je na podstawie czysto teoretycznych przesłanek, ponieważ gdyby istniała głęboka symetria między bozonami a fermionami, ich istnienie byłoby naturalną tego konsekwencją, a w fizyce symetrie zawsze odgrywają znaczącą rolę. Każdy fermion miałby zatem jednego bozonowego, a każdy bozon jednego fermionowego partnera. Na przykład do elektronu mógłby się przyłączyć partner o spinie = 0, czyli selektron. Pytanie, czy w przyrodzie istnieje supersymetria, czy też nie, należy do ważnych i wciąż otwartych pytań w świecie fizyki. Być może istnienie przeważającej części ciemnej materii we wszechświecie (zob. część 4) można przypisać supersymetrycznej cząstce. Wygląda na to, że teoria strun także nie może się obejść bez supersymetrii (część 5).

Kwarki i leptony nie są jedynymi cząstkami modelu standardowego – uzupełniają je cząstki tak zwanych pól cechowania. Cząstki te przekazują z kolei określone oddziaływania: foton – oddziaływanie elektromagnetyczne, a gluony – oddziaływanie silne. Oprócz nich istnieją jeszcze cząstki odpowiedzialne za oddziaływanie słabe, które nazywane są bozonami W i bozonami Z. Chodzi tu o bozony o spinie = 1.

Określenie pól cechowania wskazuje na fundamentalną właściwość modelu standardowego. Chodzi o teorię cechowania. Właściwość teorii cechowania polega na tym, że istnieją w nich wielkości, które dają się „przecechować", czyli które można transformować w dobrze zdefiniowany sposób bez zmiany obserwowanych wielkości. Teoria posiada wtedy „symetrię cechowania". Już elektrodynamika zna taką wielkość, tak zwany „potencjał wektorowy". Został on odkryty przez Gaussa. Zgodnie z jego dziennikiem stało się to 23 stycznia 1835 roku o siódmej rano przed wstaniem z łóżka. Obserwowalne są pola elektryczne i magnetyczne, które się nie zmieniają, gdy potencjał wektorowy jest transformowany.

Początkowo zdaje się, że zmienne cechowania są po prostu wielkościami pomocniczymi o wygodnych matematycznych właściwościach. W rzeczywistości modelu standardowego nie można

sformułować w inny sposób. Ze wszystkimi polami cechowania powiązane są bowiem cząstki odpowiedzialne za obserwowane w przyrodzie oddziaływania. Symetria cechowania to fundamentalna zasada przyrody, której źródło nie jest jeszcze w pełni zrozumiałe.

Warto wspomnieć o jeszcze jednym polu związanym z tak zwanymi cząstkami Higgsa (nazwa pochodzi od nazwiska Petera Higgsa, który postulował istnienie tego pola w 1964 roku równocześnie, ale niezależnie od Roberta Brouta i François Englerta). Nie zostało ono jeszcze odkryte, lecz ma zasadnicze znaczenie w modelu standardowym. Jest to bowiem pole, które z powodu swojego oddziaływania z innymi cząstkami jest odpowiedzialne za posiadanie przez nie masy. Pochodzenie tej masy jest w modelu standardowym bezpośrednio powiązane z cząstkami Higgsa.

Nie dziwi zatem fakt, że w akceleratorach na całym świecie trwają poszukiwania tej słynnej cząstki. Nadzieje pokładane są przede wszystkim w *Large Hadron Collider* (LHC) w CERN, czyli Europejskiej Organizacji Badań Jądrowych w Genewie, który ma zostać uruchomiony w 2009 roku (hadrony to cząstki podlegające oddziaływaniu silnemu). W urządzeniu tym protony lub jony ołowiu będą wystrzeliwane przeciwko sobie w kolistym podziemnym tunelu o długości około 27 kilometrów. W ciągu sekundy będzie mogło dojść do prawie miliarda zderzeń. Ilość energii uwolnionej w takim procesie powinna być wystarczająca, aby umożliwić powstanie cząstki Higgsa, o ile przewidywania dotyczące jej masy okażą się słuszne. W LHC będą prowadzone także poszukiwania potencjalnych cząstek supersymetrycznych, wskazówek odnośnie do wyższych wymiarów czasoprzestrzeni, jak również małych czarnych dziur.

Jak już wielokrotnie podkreślałem, pojęcie cząstki w teorii kwantowej nie jest proste. Jeżeli chodzi o kwantową teorię pola, należy najpierw ustalić, czym właściwie jest stan bez cząstek, czyli próżnia. Zazwyczaj definiuje się ją jako stan o najniższej energii. Dopóki nie jest uwzględniona grawitacja, znaczenie mają tylko względne energie, a dzieje się tak dlatego, ponieważ nie istnieje standard pomiaru dla wartości absolutnej. Z uwagi jednak na to, że wskutek równoważności masy

i energii każda forma energii wytwarza pole grawitacyjne, zostaje zdefiniowana absolutna wartość energii. Wartość ta jest powiązana z problemem stałej kosmologicznej, który przedstawimy w części 4. W tym miejscu pominiemy na razie grawitację.

Czasoprzestrzenią leżącą u podstaw szczególnej teorii względności jest wprowadzona w części 1 czasoprzestrzeń Minkowskiego (w skrócie zwana przestrzenią Minkowskiego). Stan próżni pola w przestrzeni Minkowskiego jest tak zdefiniowany, że przynależna do niego energia jest równa zeru. W wyniku nieuniknionej obecności cząstek wirtualnych próżnia nie odpowiada już tradycyjnemu wyobrażeniu o próżni jako absolutnej pustce. Z tego powodu dodawanie znikającej energii nie jest proste, a naiwne obliczenia skutkują początkowo nieskończenie wielką wartością energii, która wynika z sumy nieskończenie wielu cząstek wirtualnych. Z uwagi jednak na fakt, że tylko względne energie mają znaczenie, można tę nieskończenie wielką energię próżni „zrenormalizować" (mówiąc językiem fachowym) i ustalić na poziomie zera. Wspomniane powyżej dodawanie dotyczy w pierwszej kolejności tylko układów inercjalnych. Obserwator poruszający się ruchem jednostajnym prostoliniowym w próżni Minkowskiego nie zarejestruje zatem w swoim detektorze żadnych cząstek. Stan kwantowy próżni jest globalnym stanem, który splątuje ze sobą wszelkie obszary w przestrzeni Minkowskiego.

O tym, jak bardzo pojęcie próżni jest w stanie oszukać wyobrażenie o danej sytuacji, można się przekonać, wprowadzając do próżni Minkowskiego dla pola elektromagnetycznego dwie równoległe przewodzące (lecz nienaładowane) płyty. Można się spodziewać, że próżnia w obrębie płyt pozostanie niezmieniona, ale takie oczekiwanie jest błędne. Obie płyty przyciągają się z siłą, która jest odwrotnie proporcjonalna do czwartej potęgi odległości między płytami. Przyciąganie to jest wynikiem ujemnej energii, która znajduje się między płytami. Taki zdumiewający efekt przewidział już w 1948 roku niderlandzki fizyk Hendrik Casimir (1909–2000). Udało się go

potwierdzić eksperymentalnie z dużą dokładnością dopiero w 1997 roku, przy czym podczas testów wykorzystano przyciąganie kuli i płyty.

Dlaczego próżnia w „efekcie Casimira" posiada nagle energię między płytami, i to w dodatku ujemną? Możliwym wyjaśnieniem jest to, że w myśl teorii kwantowej z cząstkami wirtualnymi, podobnie jak z normalnymi cząstkami, łączy się pojęcie długości fali de Broglie'a. Dopóki nie wystąpią pewne ograniczenia, możliwe są wszystkie długości fali od zera do nieskończoności. W obecności płyt już nie wszystkie długości fali mogą tam wystąpić, gdyż przyjmuje się, że między płytami powstaje węzeł fali. W przestrzeni między płytami jest miejsce tylko na połowę długości fali lub jej całkowite wielokrotności. Brakuje zatem wielu długości fali i tym samym związanych z nimi energii. Skutkiem tego jest przesunięcie się energii całkowitej między płytami do wartości poniżej zera. „Nadliczbowe" drgania na zewnątrz naciskają w pewnym sensie na płyty, które wskutek tego przyciągają się. Podobny efekt znany jest w nautyce. Dwie łodzie, płynące równolegle, zbliżają się do siebie, gdy istnieją pewne profile prądów wywołane przez wiatry, a ponieważ między łodziami brakuje określonych długości fali, nadliczbowe długości fali napierają od zewnątrz.

Jeżeli chodzi o grawitację, energia według Casimira spełnia jak każda masa zasadę równoważności, czego dowiodły szczegółowe obliczenia amerykańskiego fizyka Stephena Fullinga i jego współpracowników w 2007 roku. Energia według Casimira „spada" więc w polu ciężkości tak szybko jak każda inna masa.

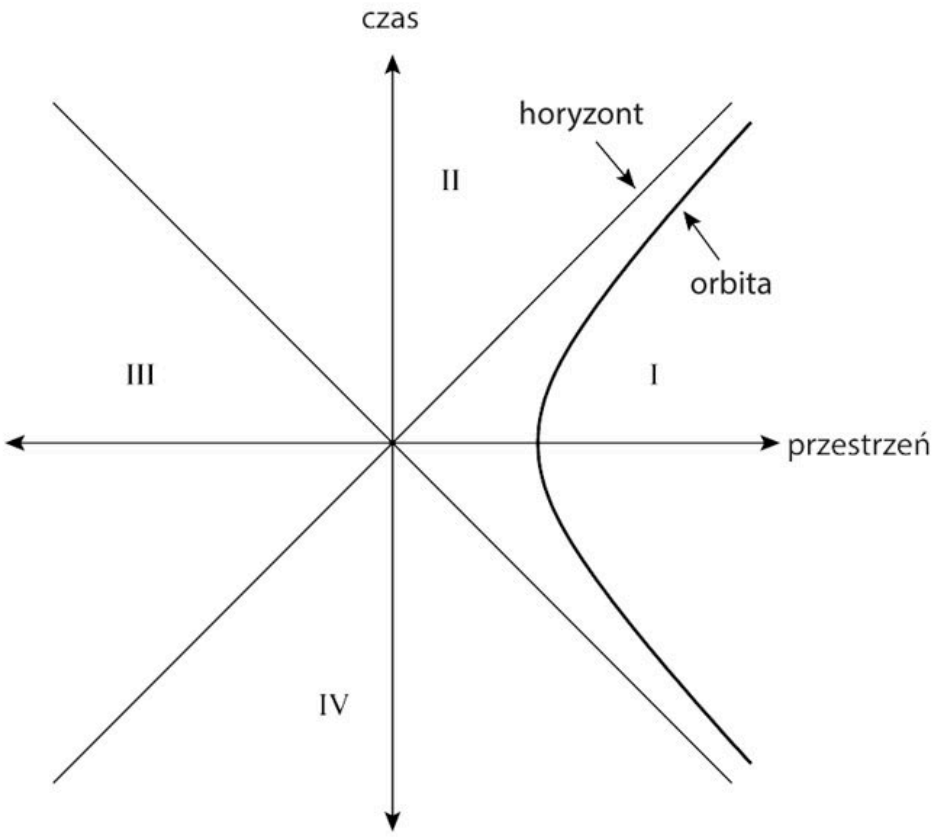
Przyciąganie obu płyt wykazuje zdumiewającą analogię do innej siły, która panuje między nienaładowanymi molekułami – siły van der Waalsa, nazwanej tak na cześć niderlandzkiego fizyka Johannesa van der Waalsa (1837–1923). Jej źródło można odnaleźć w nieuniknionych fluktuacjach rozkładu elektronów w molekułach, wynikających z teorii kwantowej. Jest to siła, która umożliwia na przykład gekonowi, egzotycznej jaszczurce, przemykanie po sufitach bez upadku. Między neutralnymi przewodnikami

siła Casimira jest siłą dominującą od wartości około 10 nanometrów w górę – poniżej tej wartości silniejsza jest siła van der Waalsa.

Próżnia Minkowskiego dla pola elektromagnetycznego (i każdego innego pola) zmienia się zatem, gdy pojawiają się płyty lub inne obiekty geometryczne. Wskutek tego zmianie ulega globalnie zdefiniowany i splątany stan próżni. Zaskakujące jest to, że zmiana taka zachodzi również poza układem inercjalnym. Obserwator, który porusza się prostoliniowo ze stałym przyspieszeniem w próżni Minkowskiego, odczuwa tę próżnię, jakby była ona wypełniona cząstkami, a jego detektory są odpowiednio pobudzone. Rozkład tych cząstek także nie jest dowolny. Odpowiada on ciepłej kąpieli z określoną temperaturą, która jest proporcjonalna do przyspieszenia i kwantu działania. Jest to „efekt Unruha", nazwany tak na cześć amerykańskiego fizyka Williama Unruha, który opisał go w 1976 roku. Surowe jajko może się więc ugotować przez samo przyspieszenie (choć byłyby potrzebne nierealnie wysokie przyspieszenia).

Efekt ten można zrozumieć jedynie wtedy, gdy uwzględni się globalną naturę próżni Minkowskiego. Obserwator poruszający się ze stałym przyspieszeniem może dostrzec z reguły tylko pewną część czasoprzestrzeni Minkowskiego. Ilustrację tego zjawiska stanowi rysunek 2.9. Oś czasu skierowana jest w górę, a kierunek przestrzeni, w której porusza się obserwator – w prawo. Na takim diagramie czasoprzestrzennym linią świata obserwatora jest hiperbola, jak wynika z równań szczególnej teorii względności. Można rozpoznać, że obserwator pozostaje ograniczony do obszaru czasoprzestrzeni I. Nie może on odbierać żadnych sygnałów z obszarów II i III. Na tym diagramie promienie świetlne poruszają się po prostych, które są nachylone pod kątem 45° w stosunku do linii poziomej. Z tego powodu nie mogą się przedostać z obszarów II i III do obszaru I. Wszystkie sygnały materialne muszą poruszać się wewnątrz tego „stożka świetlnego", czyli z nachyleniem większym niż 45° w stosunku do linii poziomej, dlatego też żaden sygnał z obszarów III i II nie może dotrzeć do

obszaru I. O dwusiecznej między obszarami I i II mówi się wtedy jako o *horyzoncie*. Horyzont ten zmienia próżnię podobnie jak płyty w efekcie Casimira.



2.9. Diagram czasoprzestrzenny dla obserwatora poruszającego się ruchem jednostajnie przyspieszonym

Dla obserwatora poruszającego się ruchem przyspieszonym częściowy obszar I jest w pewnym sensie całkowitą czasoprzestrzenią. Z tego powodu musi on zdefiniować swoją próżnię w odniesieniu do tego obszaru częściowego. Oczywiście nowa próżnia nie może mieć żadnych korelacji, czyli żadnego splątania, z obszarami wykraczającymi poza czasoprzestrzeń całkowitą dla obserwatora znajdującego się w obszarze I. Stara próżnia Minkowskiego posiada jednak te korelacje, dlatego nie może pokrywać się z tą nową. W związku z tym obserwator poruszający się ruchem

przyspieszonym widzi w próżni Minkowskiego „cząstki”. Pojęcie próżni i bazujące na nim pojęcie cząstki nie są już żadnymi niezmienniczymi pojęciami w ramach kwantowej teorii pola. Pojęcie cząstki na płaszczyźnie fundamentalnej jest bez znaczenia.

Promieniowanie czarnych dziur

Próżnia w kwantowej teorii pola jest przedziwnym tworem. To, co nazywamy próżnią, zależy od rzeczywistego stanu obserwatora. Inercjalny obserwator w przestrzeni Minkowskiego nazywa próżnią coś, co dla obserwatora poruszającego się ruchem jednostajnie przyspieszonym jest wypełnione cząstkami, które są tak rozłożone, jakby tworzyły ciepłą kąpiel z określoną temperaturą (dlatego mówi się o promieniowaniu termicznym). Głębszą przyczynę tych fenomenów stanowi nielokalna struktura próżni, czyli stan splątany obejmujący całkowitą przestrzeń Minkowskiego.

Poznaliśmy zasadę równoważności jako podstawę ogólnej teorii względności. Obserwator poruszający się ruchem przyspieszonym w przestrzeni, w której nie oddziałuje siła ciężkości, jest lokalnie równoważny z obserwatorem w polu grawitacyjnym. Czy zatem w polach grawitacyjnych nie powinno występować też coś na kształt omówionego powyżej efektu Unruha? Rzeczywiście tak jest. Istotna w efekcie Unruha jest obecność horyzontu – obserwator poruszający się ruchem przyspieszonym może dostrzec z reguły tylko pewną część czasoprzestrzeni Minkowskiego. Podobna sytuacja zachodzi w przypadku czarnych dziur. Obiekty te składają się z czystej czasoprzestrzeni, a powstają wskutek kolapsu grawitacyjnego masywnych gwiazd, po tym jak ich zasoby energetyczne ulegną wyczerpaniu (zob. część 1). Jedną z głównych właściwości czarnej dziury jest istnienie horyzontu. Teoria Einsteina zakłada, że żaden sygnał z obszaru wewnątrz horyzontu nie może wydostać się na zewnątrz. Obliczenia potwierdzają już, że struktura czasoprzestrzeni w pobliżu horyzontu czarnej dziury jest podobna do otoczenia horyzontu u obserwatora z teorii Unruha. Z tego względu

również w czarnej dziurze należy spodziewać się promieniowania termicznego.

Skąd miałoby pochodzić to promieniowanie? W 1974 roku, w jednym z najbardziej fascynujących opracowań fizyki teoretycznej ostatnich dziesięcioleci, Stephen Hawking wykazał, że czarne dziury emitują promieniowanie, gdy do ich opisu włączy się teorię kwantową. Promieniowanie to jest określone przez temperaturę proporcjonalną do kwantu działania, nazwaną „temperaturą Hawkinga". Hawking podał zrozumiałe wyjaśnienie tego efektu i źródło promieniowania. Próżnię przenikają cząstki wirtualne. Zgodnie z tym, co przedstawiono na rysunku 2.8, foton może zamienić się w parę elektron-pozyton i powrócić do pierwotnego stanu. W normalnych okolicznościach para taka jest nietrwała i poza tym nic więcej się nie dzieje. W otoczeniu horyzontu czarnej dziury grawitacyjne siły pływowe mogą być tak silne, że prowadzą do rozerwania tej wirtualnej pary. Jedna z cząstek wpada do czarnej dziury, drugiej zaś może udać się ucieczka na zewnątrz. Pole grawitacyjne czarnej dziury umożliwia cząstkom wirtualnym przeobrażenie się w cząstki realne, które można zaobserwować. Z powodu uciekających cząstek par z zewnątrz wygląda to tak, jakby czarna dziura faktycznie emitowała promieniowanie. Z uwagi na to, że wszystkie istniejące w przyrodzie cząstki w swojej wersji wirtualnej poruszają się w próżni, zostaną one również w komplecie wyemitowane przez czarną dziurę. W zależności od temperatury jest to jednak możliwe tylko w przypadku tych cząstek, których masa i przez to energia spoczynkowa są odpowiednio niewielkie, aby umożliwić emisję energii termicznej.

Oczywiście taki obraz cząstek jest jak zwykle niezbyt precyzyjny, lecz znajduje swoje odzwierciedlenie w dokładnych wzorach z dziedziny teorii kwantowej. Z obliczeń Hawkinga wynika przepływ ujemnej energii, która wpływa do wnętrza czarnej dziury. Ta ujemna energia przypomina ujemną energię w efekcie Casimira i jest nieuniknioną konsekwencją teorii kwantowej. Gdy wpływa ona do dziury, odbiera jej dodatnią energię. Jedyną energię, jaką posiada niewirująca dziura, jest energia spoczynkowa,

która tkwi w jej masie. W konsekwencji promieniowanie Hawkinga odbiera jej masę – dziura staje się lżejsza wskutek emisji promieniowania. Z uwagi na to, że energia zostaje zachowana, strumień dodatniej energii, która tkwi w wyemitowanym promieniowaniu, wypływa na zewnątrz.

Ze względu na to, że proces zachodzi w bezpośrednim zewnętrznym otoczeniu horyzontu, w gruncie rzeczy żadne promieniowanie nie wydostaje się z czarnej dziury na zewnątrz. Oddalony obserwator ma jednak takie wrażenie. Obserwator, który spada swobodnie przez horyzont do wnętrza dziury, nie dostrzega natomiast żadnego promieniowania. To również jest skutkiem zasady równoważności. Swobodnie spadający obserwator odpowiada inercjalnemu obserwatorowi, który także nie dostrzega promieniowania w zwykłej próżni Minkowskiego.

W niewirującej czarnej dziurze (według teorii Schwarzschilda) temperatura Hawkinga zależy, poza stałymi fizycznymi, wyłącznie od jej masy. Temperatura jest jednak odwrotnie proporcjonalna do tej masy. Oznacza to, że przy malejącej masie (tak dzieje się wskutek emitowania promieniowania Hawkinga) temperatura wzrasta. Czarna dziura staje się bardziej *gorąca* w wyniku emisji promieniowania. Jest to dokładnie odwrotne zachowanie w stosunku do tego, które cechuje zwykłe układy termodynamiczne. Gorąca herbata schładza się przez emisję promieniowania, a nie staje się przez to bardziej gorąca. „Zwariowane" termodynamiczne zachowanie czarnych dziur jest skutkiem grawitacji i jej uniwersalnej przyciągającej natury. Obiekt zyskuje zatem energię, gdy w polu grawitacyjnym zbliża się do przyciągającej masy, z czym mamy do czynienia także w przypadku kurczącej się w wyniku emisji promieniowania czarnej dziury.

Wspomniane stałe fizyczne, które mają wpływ na temperaturę Hawkinga, to stała grawitacyjna, prędkość światła i kwant działania (jak również stała termodynamiczna, którą poznamy w następnej części tej książki). Wskazuje to na fundamentalną teorię, w której efekty teorii względności i teorii kwantowej są w równym stopniu istotne i przez to obie są w tej teorii uwzględnione. Niestety proces opracowywania takiej

teorii, „teorii grawitacji kwantowej", nie jest jeszcze zakończony. W części 5 zobaczymy, które elementy tej teorii są już znane oraz co z nich wynika dla zrozumienia kwantowego kosmosu.

Skoro czarna dziura wskutek emisji promieniowania Hawkinga stale traci na masie, to ostatecznie musi dotrzeć do punktu, w którym całkowita masa zostanie zużyta. Co się wtedy stanie? W tej końcowej fazie teoria Hawkinga traci swoją skuteczność. Należałoby posłużyć się niekompletną teorią grawitacji kwantowej, aby zrozumieć ostatni etap parowania. O tym jednak później. Długość życia czarnej dziury aż do jej końcowej fazy można mimo wszystko oszacować. Dla czarnej dziury o masie Słońca wynosi ona niewyobrażalne 10^{65} lat, co stanowi wartość o wiele potęg liczby 10 większą niż wiek całego wszechświata (część 4). Im lżejsza jest czarna dziura, tym krócej żyje. Aby otrzymać wartość długości życia porównywalną z obecnym wiekiem wszechświata, wynoszącym około 10 miliardów lat, potrzebowalibyśmy dziury o masie początkowej rzędu 500 milionów ton. Tyle wynosi masa niewielkiej asteroidy. Taka względnie lekka czarna dziura jest jednak mikroskopijna. Osiąga ona bowiem ledwie średnicę jądra atomu! W ostatniej sekundzie swojej egzystencji wyemituje około 10^{23} dżula energii.

Wartość temperatury Hawkinga dla czarnych dziur o kilku masach Słońca i większych jest niemierzalnie niska. Dziura o masie Słońca posiadałaby temperaturę wynoszącą zaledwie około 10 milionowych kelwina, a zatem zbyt niską, aby móc ją zmierzyć metodami stosowanymi w astrofizyce.

Wyłącznie czarne dziury o masie w wyżej wymienionym zakresie około 500 milionów ton i mniejszych posiadają wystarczająco wysoką temperaturę, aby mogło dojść do zdarzenia pozwalającego dostrzec je za pomocą teleskopu. Czy w ogóle istnieją czarne dziury o odpowiednio małej masie? Na pierwszy rzut oka wydaje się to niemożliwe. Zgodnie z teorią rozwoju gwiazd czarne dziury, które powstają jako końcowe stadium gwiazd, muszą mieć masę odpowiadającą co najmniej trzem masom Słońca. Jest jednak jeszcze możliwość, że czarne dziury powstały

we wczesnej fazie wszechświata, kilka sekund po wielkim wybuchu w wyniku dużych fluktuacji gęstości. W przypadku tak zwanych pierwotnych czarnych dziur nie istnieje żadne ograniczenie co do masy. Niestety nie dysponujemy wystarczającą wiedzą na temat zjawisk z początku wszechświata, aby móc dokładnie obliczyć szybkość powstawania tych obiektów. Z pewnością nie powstało ich jednak zbyt wiele, ponieważ w przeciwnym razie obecnie powinno ich być tak dużo, że nie mielibyśmy problemu z ich dostrzeżeniem, a jak dotąd mimo wielu starań nie udało nam się zaobserwować choćby jednej. W jaki sposób moglibyśmy je zidentyfikować? Temperatura Hawkinga w ostatniej fazie parowania czarnej dziury powinna być tak duża, że dochodziłoby wtedy przede wszystkim do emisji promieniowania gamma o charakterze termicznym. Z uwagi na to, że zjawiska takiego nie udało się jeszcze zaobserwować, można podać jedynie górne granice. W astrofizyce można co prawda odnotować rozbłyski gamma, które są jednak tak wysokoenergetyczne, że ich źródłem nie mogą być czarne dziury. Badano też, czy promieniowanie czarnych dziur Hawkinga może przyczynić się do wzbogacenia ogólnej wiedzy o promieniach gamma, lecz i w tym przypadku można podać jedynie górne granice. Być może w kolejnym pokoleniu akceleratorów takich jak LHC będzie można bezpośrednio wytwarzać i badać czarne dziury (zob. także część 5). Dopóki nie będzie można zaobserwować takich obiektów, jeden z najbardziej fascynujących fizycznych efektów nie zostanie potwierdzony eksperymentalnie.

3. Entropia

Czas jest czynny, ma takie właściwości, jak czasowniki. Ma „poczesny” udział w tworzeniu.
Ale czego? Zmiany!

Thomas Mann, *Czarodziejska góra*
(wg przekładu Jana Łukowskiego, 1965)

Druga zasada termodynamiki

Gorąca herbata w filiżance schładza się, jednocześnie ogrzewając otoczenie. Jak dotąd nikt nie zauważył, aby ciepło otaczającego nas powietrza napływało do gorącej herbaty i ją podgrzewało. W układach pozostawionych samym sobie ciepło płynie zawsze od gorętszego do chłodniejszego obiektu, nigdy odwrotnie. Wiele fenomenów w przyrodzie wykazuje taką nieodwracalność (bezpowrotność). Spadająca na podłogę szklanka tłucze się, a stłuczone kawałki nigdy nie połączą się z powrotem w szklankę. Owoce gniją i nigdy nie staną się ponownie świeże. Nieodwracalność przenika całe życie, które nieuchronnie kończy się śmiercią. Istnieje różnica między przeszłością a przyszłością, która zdaje się jednak nie mieć żadnego znaczenia w podstawowych równaniach fizyki. Zasady mechaniki Newtona znają wprawdzie pojęcie czasu absolutnego, lecz nie znają absolutnego kierunku tego czasu. Gdy wyobrazimy sobie odwrócony ruch planet, jak w filmie, który przewijamy do tyłu, otrzymamy takie samo rozwiązanie równań Newtona.

Zagadnienie nieodwracalności wzbudziło ogromne zainteresowanie w świecie fizyki dopiero w XIX wieku. Wyraża się ono przede wszystkim w pojęciu entropii. Jest to jedno z najbardziej fundamentalnych i zagadkowych pojęć naukowych, które odnosi się także do innych dyscyplin, takich jak teoria sztuki czy ekonomia. Nieco potocznie entropia

rozumiana jest jako miara nieporządku. Co naprawdę kryje się za tym pojęciem? Jakie znaczenie ma pojęcie entropii dla kwantowego kosmosu?

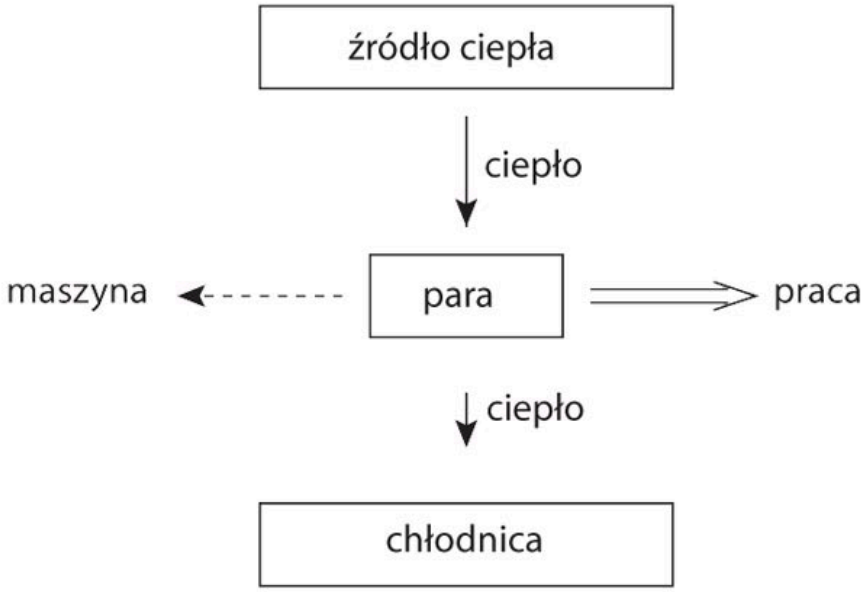
Słowo pochodzi od greckiego *tropé*, co oznacza przemianę. Jeden z podręczników do fizyki wyjaśnia zatem entropię jako wewnętrzną zdolność zamkniętego układu do zmiany swojego stanu. W tym zdaniu została wyrażona jedna z jej głównych właściwości, a mianowicie entropia zmienia się także dla układu całkowicie pozostawionego sobie, a dokładniej na ogół rośnie. Słowo to nawiązuje do pojęcia energii, wprowadzonego do fizyki w dzisiejszym jego rozumieniu mniej więcej w tym samym czasie. Energia jest jednak stała w układzie zamkniętym, podczas gdy entropia rośnie lub pozostaje taka sama.

Pojęcie entropii zostało wyodrębnione z technicznych rozwiązań z początków epoki przemysłowej. Symbol tych czasów, maszyna parowa, miał w tym przypadku decydujące znaczenie. Dzięki niej zamieniano ciepło w pracę, a praca miała fundamentalne znaczenie dla potrzeb nowej epoki: transportu i komunikacji. Pojęcie to wprowadził niemiecki fizyk Rudolf Clausius (1822–1888), który w 1865 roku napisał: „Energia wszechświata jest stała. Entropia wszechświata dąży do maksimum" (wg przekładu Andrzeja Kajetana Wróblewskiego, 2006). Z jakiego powodu entropia rośnie i dlaczego ma to coś wspólnego z obserwowaną nieodwracalnością świata?

Clausius oparł swoje przemyślenia na pracach francuskiego fizyka Sadiego Carnota (1796–1832), który podczas wizyty u swojego ojca będącego na wygnaniu w Magdeburgu zajął się problemami maszyny parowej. Odkrył on, że ciepła nie można całkowicie zamienić w pracę. Carnot zyskał to słuszne przeświadczenie jeszcze na podstawie przestarzałego wyobrażenia, że ciepło jest związane z własną substancją cieplną, tak zwanym cieplikiem, która podczas podgrzewania wstępuje w ciało, a podczas schładzania je opuszcza. Słowo *calorique*, którego używano w związku z cieplikiem, przetrwało w związanych z żywieniem kaloriach, które spożywamy (bądź których unikamy).

Główna publikacja Carnota z 1824 roku początkowo przeszła bez echa. Nie bez znaczenia był tu fakt, że pojęcia energii i zachowania energii w związku z ciepłem były wtedy jeszcze niezrozumiałe. Nie było jasne, która z wielkości podczas rozważanych procesów zostaje zachowana, a która nie. Do wyjaśnienia pojęcia energii jako pierwszy przyczynił się lekarz z Heilbronnu Julius Robert von Mayer (1814–1878), który odkrył, że ciepło jest tylko pewną formą energii i że suma wszystkich energii, która składa się między innymi z ciepła i pracy, jest wielkością zachowaną. Podał on także miarę ilościową równoważności pracy i ciepła. Gdy na przykład masa 42,7 kilograma spada z wysokości 10 metrów, to ilość wytworzonej w ten sposób energii wystarcza na ogrzanie jednego litra wody o jeden stopień. Niezależnie od Mayera z dużą dokładnością wyznaczył ten równoważnik ciepła brytyjski fizyk James Prescott Joule (1818–1889), od którego nazwiska pochodzi nazwa jednostki energii, dżula (jedna kaloria to nieco ponad cztery dżule). Co ciekawe, Mayer doszedł do swoich wyników podczas studiowania procesów przemiany materii w organizmie, gdzie ciepło jest wytwarzane poprzez spalanie substancji odżywczych w sposób bezpośredni bądź pośredni przez mechanikę mięśni. Zachowanie całkowitej energii w związku z procesami cieplnymi nazwano później pierwszą zasadą termodynamiki (nauki o cieple).

Dopiero gdy pojęcie energii zostało wyjaśnione, a ciepło i praca okazały się wzajemnie przekształcającymi się w siebie formami energii, zaczęto zastanawiać się nad wynikami Carnota. Dlaczego praca całkowicie zmienia się w ciepło, a ciepło nie pozwala się całkowicie przekształcić w pracę? Skąd bierze się związana z tym procesem nieodwracalność? Warto dokładniej przyjrzeć się całemu procesowi, nad którym pracował Carnot. Wprawdzie chodzi tu o wyidealizowany przypadek graniczny, lecz jest on bardzo pouczający. Proces Carnota opisuje maszynę pracującą cyklicznie, czyli zawsze powracającą do punktu wyjścia. Inne maszyny nie sprawdzają się w praktyce, ponieważ na przykład w przypadku maszyny parowej celem jest nie tylko jednorazowe, ale regularne pozyskiwanie pracy. Cykl dzieli się na cztery procesy (porównaj z rys. 3.1).



3.1. Cykl Carnota

W pierwszym procesie ciepło wypływa ze źródła ciepła w formie pary, która się ogrzewa, rozszerza i przy tym jest w stanie wytworzyć pracę, na przykład dzięki temu, że zostaje poruszony tłok. Gdybyśmy chcieli w tym miejscu odwrócić proces, aby powrócić do stanu wyjściowego, nic byśmy nie osiągnęli. Należałoby włożyć taką samą pracę, jaką właśnie się uzyskało. Z tego powodu dokłada się kolejne trzy procesy. W drugim procesie ciśnienie gazu zostaje zmniejszone, co osiąga się przez ekspansję po odcięciu źródła ciepła (tego rodzaju proces określany jest jako adiabatyczny). W tym procesie para wykonuje jeszcze pracę. W trzecim procesie objętość pary zmniejsza się ponownie, gdy para jest w kontakcie z chłodnicą i przekazuje do niej swoje ciepło. Temperatura chłodnicy powinna być tak niska, jak to tylko możliwe, aby zminimalizować niezbędny wkład pracy. Ostatecznie w czwartym procesie gaz spręża się adiabatycznie do swojej poprzedniej wartości. Maszyna ponownie znajduje się w stanie wyjściowym, a proces może rozpocząć się od nowa.

Co uzyskano w procesie Carnota? Ciepło zostało pozyskane ze źródła ciepła, a mniejsza jego ilość musiała zostać przekazana do chłodnicy. Różnica (pozyskane minus przekazane ciepło) może zostać zamieniona w pracę, co jest wynikiem ruchu tłoka. Dokładne zbadanie tego cyklicznego procesu pokazuje, że ciepło nigdy nie może być całkowicie zmienione w pracę. Można zbliżyć się do tego ideału, gdy źródło ciepła będzie miało możliwie najwyższą temperaturę, a chłodnica możliwie najniższą. Różnice temperatury mają więc duże znaczenie. W procesie Carnota przy odpowiednich źródłach ciepła można osiągnąć najwyższą możliwą wydajność pracy. Realna maszyna ma niższą wydajność. Proces Carnota wyznacza teoretyczną górną granicę, która jest niezależna od materiału i sztuki inżynierskiej. Wprowadzona później przez Clausiusa entropia w procesie Carnota zostaje zachowana, a w każdym innym procesie rośnie. Przepływ ciepła w kierunku od gorąca do zimna we wszechświecie zwiększa zatem entropię wszechświata.

Pierwsza zasada termodynamiki głosi, że energia układu zamkniętego zostaje zachowana. Natomiast według drugiej zasady termodynamiki entropia układu zamkniętego nigdy nie może się zmniejszyć. Może być stała w wyidealizowanych przypadkach granicznych, z reguły jednak rośnie. Clausius wyraził to w cytowanym już wcześniej zdaniu, że entropia wszechświata dąży do maksimum.

Drugą zasadę można sformułować także na wiele innych sposobów, które są nieco bardziej obrazowe. Sam Clausius ujmował brak ubytku entropii w taki sposób, że nie ma takiego procesu, w którym jedynym wynikiem jest przekazanie energii od ciała chłodniejszego do cieplejszego. Ciepło samo z siebie płynie zawsze w kierunku od cieplejszego do chłodniejszego obiektu. Brytyjski fizyk William Thomson (1824–1907), późniejszy lord Kelvin, sformułował zasadę w wyżej wymienionej formie, że całkowita zamiana ciepła w pracę nie jest możliwa. Mówi się również o niemożności skonstruowania *perpetuum mobil*e drugiego rodzaju (*perpetuum mobile* pierwszego rodzaju naruszyłoby zasadę zachowania energii, ponieważ mogłoby wytwarzać energię z niczego). Z tego powodu

nie jest możliwe rozwiązanie problemu energetycznego ludzkości, wykorzystując w tym celu gigantyczną energię cieplną tkwiącą w oceanach i atmosferze, aby zamienić ją w pracę.

Można udowodnić równoważność różnych wersji drugiej zasady termodynamiki. Wszystkie wyrażają fundamentalny aspekt jej natury, czyli nieodwracalność. Istnieje wiele procesów, które mogą przebiegać w czasie tylko w określonym kierunku, a wzrost entropii wyraża to oznaczenie kierunku czasu. Ciepło jest bardziej nieuporządkowaną formą energii niż praca, dlatego też z biegiem czasu nieporządek przybiera na sile, jak głosi obrazowa interpretacja entropii.

W układzie, który nie jest zamknięty (takie układy nazywane są *otwartymi*), entropia może się zmniejszać. Płynie ona wtedy z układu do otoczenia, wskutek czego całkowita entropia układu i otoczenia rośnie. Najprostszym przykładem jest lodówka. Ciepło pompowane jest ze środka na zewnątrz i otoczenie ogrzewa się. Entropia w środku jest sztucznie podtrzymywana na niskim poziomie, podczas gdy entropia otoczenia rośnie. Nie inaczej funkcjonuje fundamentalna zasada każdego życia. Także w przypadku żywej istoty entropia jest utrzymywana na niskim poziomie, ponieważ jest oddawana do otoczenia. Jest to możliwe tylko dlatego, że w kosmosie mamy do czynienia z ogromnymi różnicami temperatur, na przykład gorące Słońce pośród zimnej przestrzeni kosmicznej. Z tego powodu możemy naszą entropię utrzymać na niskim poziomie, oddając ją na zewnątrz. Przyjmujemy pożywienie w formie uporządkowanej o niskiej entropii (którego istnienie jest zagwarantowane jedynie przez energię słoneczną) i oddajemy energię o wysokiej entropii w formie potu i wydzielin. Takie sztuczne utrzymywanie entropii na niskim poziomie nie funkcjonuje jednak wiecznie, każde życie kończy się przecież śmiercią. W tym miejscu widać wyraźnie, że druga zasada musi mieć swoje kosmiczne źródło. Wszechświat bez różnic temperatury osiągnąłby już maksymalną entropię i życie nie mogłoby istnieć. W XIX wieku określano taki stan jako śmierć cieplną, która nieuchronnie grozi wszechświatowi.

Czy istnieje jakaś głębsza przyczyna obowiązywania drugiej zasady, czy po prostu musimy ją uznać za fakt empiryczny? Głębsze uzasadnienie istnieje, gdy uwzględnimy atomową strukturę materii. Uczynimy to w kolejnym rozdziale. Teraz przyjrzyjmy się jeszcze krótko dwóm pozostałym zasadom termodynamiki. W tak zwanej zerowej zawarte jest twierdzenie, że ciała, które są ze sobą w równowadze, wykazują wspólną temperaturę – nastąpiło zatem wyrównanie ciepła. Istnieje jeszcze trzecia zasada sformułowana przez Walthera Nernsta (1864–1941) w 1905 roku, za którą w 1920 roku przyznano mu Nagrodę Nobla w dziedzinie chemii. Głosi ona, że niemożliwe jest osiągnięcie zera absolutnego temperatury w skończonej liczbie kroków. Planck ujął później tę zasadę w taki sposób, że entropia podczas zbliżania się do punktu zerowego musiałaby osiągnąć zero lub inną niezależną od materii stałą. Głębsze zrozumienie trzeciej zasady dostarcza dopiero uwzględnienie teorii kwantowej.

Mikroskopowa interpretacja entropii

Zasady termodynamiki przez długi czas były postrzegane jako fakty empiryczne, których nie można głębiej uzasadnić. Uległo to zmianie dopiero wtedy, gdy kwestią tą zainteresował się austriacki fizyk Ludwig Boltzmann (1844–1906). Boltzmann wierzył, że rozkład materii nie jest nieprzerwany, lecz składa się ona z dyskretnych jednostek, czyli atomów i molekuł. W związku z tym w termodynamice można odwołać się do mechaniki poszczególnych atomów i molekuł („cząstek”). Całkowitą energią gazu jest wtedy suma wszystkich pojedynczych energii, a jego temperaturę można zdefiniować bezpośrednio na podstawie średniej energii cząstki – im większa średnia energia, tym wyższa temperatura.

A co się dzieje z entropią? Przy próbie zrozumienia tego pojęcia po raz pierwszy do świata fizyki wkroczyły teorie dotyczące prawdopodobieństwa, zajmując w niej centralne miejsce. Stało się to ponad 50 lat przed powstaniem teorii kwantowej. W słynnej pracy z 1872 roku Boltzmann wprowadził pojęcie entropii jako miary dla liczby możliwych mikrostanów, które może posiadać gaz przy danych

wielkościach makroskopowych, takich jak temperatura, objętość, ciśnienie lub energia. Pod pojęciem mikrostanów rozumie się możliwe rozkłady cząstek pod względem ich położenia i prędkości. Jeżeli w danym makrostanie istnieje wiele możliwości, stan ten jest prawdopodobny. Jeżeli istnieje niewiele możliwości, jest on mało prawdopodobny. W rzeczywistości jest tak wiele cząstek, że podążanie za jedną z nich jest niemożliwe. W pudełku o kilkucentymetrowej długości krawędzi znajduje się mniej więcej 10^{23} cząstek. Z tego powodu można zmierzyć jedynie wielkości makroskopowe, takie jak temperatura i objętość. Entropia jest niewielka, gdy możliwa jest niewielka ilość mikrostanów, jak w przypadku molekuł gazu stłoczonych na niewielkiej przestrzeni. Entropia jest duża, gdy molekuły mają do dyspozycji całe pudełko, ponieważ cząstki mają wtedy więcej możliwości wyboru położenia. Z uwagi na to, że cząstki stale zderzają się ze sobą, istnieje prawdopodobieństwo rozkładu nie tylko pod kątem miejsca, ale i prędkości.

Weźmy prosty przykład. Pudełko jest podzielone na dwie równe części przesuwną ścianką. Na początku gaz znajduje się w jednej z dwóch części. Gdy ścianka zostanie nieco przesunięta, tworząc szczelinę, gaz będzie się stopniowo przedostawał do drugiej części pudełka, dopóki w obu częściach pudełka nie utworzy się ten sam stan. Przyczyna tego jest prosta i polega na tym, że cząstki ze względu na zderzenia zarówno między sobą, jak i ściankami pudełka podlegają przypadkowym ruchom i chcą zająć całą przestrzeń, jaką mają do dyspozycji. Gdy taki stan zostanie osiągnięty, mówimy o równowadze. Cząstki nadal zderzają się ze sobą, ale zderzenia te są właściwie takie same i odpowiadają w pewnym sensie tylko przesortowaniu cząstek.

Oczywiście istnieje także możliwość, że cząstki z powodu ruchu przypadkowego powrócą w całości do pierwszej części pudełka. Prawdopodobieństwo wystąpienia takiej możliwości jest tak niewielkie, że nigdy nie będzie można zaobserwować takiego procesu. Można to łatwo wytłumaczyć. Załóżmy, że istnieje n molekuł. Dla każdej molekuły prawdopodobieństwo znalezienia się w jednej z obu części pudełka (lewej

lub prawej) wynosi zawsze jedna druga. Teoretycznie odpowiada to rzutowi monetą z prawdopodobieństwem jednej drugiej dla orła i reszki. Całkowicie przypadkowemu wspólnemu ruchowi wszystkich molekuł z lewej części do prawej odpowiada prawdopodobieństwo wyrzucenia *n* razy pod rząd reszki. Już dla dziesięciu cząstek prawdopodobieństwo wynosi mniej niż jedną tysięczną. W przypadku wspomnianych 10^{23} molekuł znikomość prawdopodobieństwa wykracza poza granice naszej wyobraźni. Nawet prawdopodobieństwo skreślenia co tydzień przez całe życie sześciu poprawnych cyfr w totolotka jest o kilka rzędów wielkości wyższe niż powyżej opisane prawdopodobieństwo.

Jeżeli ktoś choć trochę zna się na komputerach, na pewno słyszał o bitach i bajtach. W jednym bicie informacji mamy do czynienia z jedną z dwóch możliwości z pary wartości tak lub nie. Jest to na przykład informacja, czy podczas rzutu monetą wypadnie orzeł czy reszka. Informacja, czy określona molekuła znajdzie się po lewej czy po prawej stronie, także odpowiada jednemu bitowi. Liczba bitów jest równa liczbie pytań z możliwością tak lub nie, które musimy zadać, aby stwierdzić, w jakim stanie znajduje się układ. W przypadku *n* monet należy zadać *n* takich pytań, tak samo w przypadku *n* molekuł i obu części pudełka. Entropię układu określa się w zasadzie (poza jednym współczynnikiem) jako liczbę bitów. W przypadku, gdy gaz na początku ograniczył się do lewej części pudełka, nie trzeba już zadawać kolejnych pytań, a entropia w odniesieniu do pytania o lewą czy prawą stronę jest równa zeru. Gdy molekuły wypełniają całą przestrzeń, liczba pytań jest równa liczbie molekuł, ponieważ trzeba zapytać o położenie każdej molekuły. Entropia jest wtedy równa *n* i zwiększyła się o *n* bitów. Entropię można byłoby więc również zidentyfikować bezpośrednio na podstawie liczby stanów. Z uwagi jednak na to, że ich ilość jest ogromna, zdecydowano się na liczbę bitów (z punktu widzenia matematyki jest to logarytm liczby stanów). Ma to i tę zaletę, że entropia jest proporcjonalna do liczby cząstek, czyli przy podwójnej ich liczbie jest też dwa razy większa. W ten sposób łatwiej jest przeprowadzać obliczenia.

Entropia jest zatem niczym innym jak miarą prawdopodobieństwa stanu. Jeżeli stan opisany parametrami makroskopowymi, takimi jak objętość i temperatura, można zrealizować w wielu mikrostanach, entropia jest wysoka; jeżeli udaje się to tylko w niewielu mikrostanach, entropia jest niska. Odkrycie Boltzmanna było tak nowatorskie, że spotkało się po części z gwałtownym sprzeciwem. Zwłaszcza Planck nie chciał zaakceptować takich statystycznych przemyśleń. Jak na ironię w 1900 roku musiał przyznać się do błędu, gdy przy użyciu statystycznych metod w duchu Boltzmanna wyprowadził on swoje słynne prawo promieniowania.

Z metody przeliczeniowej Boltzmanna dla doskonałego gazu bierze się stary wynik Clausiusa. Dzięki temu można także ustalić jeszcze nieokreślony współczynnik proporcjonalności – wielkość nazwaną na cześć Boltzmanna stałą Boltzmanna o symbolu k_B.

Statystyczny opis gazów wymaga wprowadzenia pojęcia zespołu statystycznego. Zespół ten obejmuje zadany rzeczywisty układ będący częścią dużej liczby abstrakcyjnych układów tego samego rodzaju. Wszystkie te układy posiadają te same parametry makroskopowe, takie jak temperatura i objętość, lecz różnią się mikrostanem. Liczbę wymyślonych układów dopasowuje się odpowiednio do danej częstości mikrostanów w taki sposób, aby stany prawdopodobne były reprezentowane w abstrakcyjnych układach częściej, a mniej prawdopodobne – rzadziej. Prawdopodobieństwo można wtedy precyzyjnie wyliczyć jako względną częstość.

W jaki sposób druga zasada termodynamiki wynika z mikroskopowej interpretacji entropii? Z uwagi na to, że entropia wiąże się bezpośrednio z prawdopodobieństwem stanu, wzrostowi entropii odpowiada przejście ze stanu nieprawdopodobnego do prawdopodobnego. Czy nie takie powinny być bezpośrednie oczekiwania? Owo przejście w powyższym przykładzie odpowiada po prostu przepływowi gazu, początkowo ograniczonemu do fragmentu przestrzeni, do całej przestrzeni. Boltzmannowi udało się udowodnić, że entropia zdefiniowana na poziomie

mikroskopowym rośnie z czasem. Dzieje się tak jednak przy założeniu, że w sensie statystycznym molekuły po zderzeniu można postrzegać jako niezależne od siebie (w rzeczywistości takie nie są). Korelacje w odniesieniu do przyszłości są zatem pomijane, a skutkiem tego jest nieodwracalność. Pomijanie to określa się również mianem założenia o molekularnym chaosie Boltzmanna.

Amerykański fizyk Josiah Willard Gibbs (1839–1903) zilustrował ten proces w swojej słynnej analogii atramentowej. Obserwujemy szklankę z wodą, do której wpada kropla atramentu. Gdy zamiesza się wodę, atrament rozdziela się według obserwatora (na oko) równomiernie. Odpowiada to przejściu ze stanu nieprawdopodobnego (tylko zlokalizowana kropla) do prawdopodobnego (atrament rozprowadzony wszędzie). Według Gibbsa takie wrażenie powstaje tylko podczas pobieżnej obserwacji bez uwzględnienia mikrostanu. Przynajmniej w przypadku atramentu wyidealizowanego jego przyjęta objętość nie zmieni się. Mikrostan składa się z maleńkich nitek atramentu, które wprawdzie są rozmieszczone w całej szklance wody, lecz mimo to posiadają tylko pierwotną objętość jednej kropli atramentu. Zlekceważenie tej struktury nitek odpowiada w przypadku gazu założeniu o molekularnym chaosie Boltzmanna – zostają pominięte korelacje, które nie mają skutków makroskopowych. Na określenie takiego pominięcia przyjęło się pojęcie „gruboziarnistości", ponieważ zgodnie z analogią atramentową Gibbsa oznacza ono „brak dokładnej obserwacji". Znajomość mikrostanu odpowiada oczywiście „drobnoziarnistości".

Wyprowadzenie drugiej zasady przez Boltzmanna nie przeszło bez echa. Wiedeński kolega Boltzmanna, Josef Loschmidt (1821–1895), sformułował tak zwany paradoks odwracalności. Atomy i molekuły w gazie poruszają się przecież, jak twierdził Loschmidt, zgodnie z zasadami mechaniki klasycznej. Zasady te są niezmienne w odwróconym czasie – nie odróżniają przeszłości od przyszłości. Jeżeli więc istnieją procesy prowadzące od stanu nieprawdopodobnego do prawdopodobnego, to musiałoby istnieć tyle samo procesów, które prowadziłyby od stanu

prawdopodobnego do nieprawdopodobnego. Entropia powinna zatem równie często rosnąć, jak i maleć, co jest sprzeczne z obserwacją.

W odpowiedzi Boltzmann sprecyzował swoją wypowiedź, wskazując na to, że w rzeczywistości nie mogłoby dojść do spełnienia warunków początkowych, niezbędnych do spadku entropii. W tym celu należałoby uwzględnić bardzo subtelne korelacje między poszczególnymi molekułami, aby zagwarantować, że rozprzestrzenione w całym pudełku cząstki same powrócą do jednej jego części. Obserwowany zazwyczaj wzrost entropii ma związek z tym, że tak subtelnie zadane korelacje zazwyczaj nie występują w przyrodzie.

Wyłącznie na podstawie rozważań nad kwestią prawdopodobieństwa należałoby założyć, że układ znajduje się w równowadze praktycznie zawsze, ponieważ istnieje maksymalna liczba mikrostanów, która spełnia ten warunek. Dozwolone mogą być jedynie nieznaczne wahania równowagi, przy czym prawdopodobieństwo takiego wahania (spadek i ponowny wzrost entropii) zmniejsza się wraz z jego wielkością.

Kolejne zastrzeżenie do wyprowadzonej przez Boltzmanna drugiej zasady zgłosił Ernst Zermelo (1871–1953), asystent Plancka w Berlinie, który także zasłynął dzięki fundamentalnym pracom matematycznym. Należał on tak jak Planck do grupy „energetyków", którzy uważali, że pojęcia takie jak energia i entropia są pojęciami fundamentalnymi, niewymagającymi żadnego wyjaśnienia mikroskopowego. „Paradoks powrotu", sformułowany przez Zermelo w 1896 roku, opiera się na matematycznym twierdzeniu francuskiego fizyka i matematyka Henriego Poincarégo, według którego wszystkie układy ograniczone przestrzennie z biegiem czasu dowolnie zbliżają się ponownie do każdego przyjętego kiedyś stanu. Zgodnie z tym twierdzeniem entropia nie może zatem stale rosnąć, lecz kiedyś musi się zmniejszyć, aby móc ponownie osiągnąć ewentualny stan wyjściowy niskiej entropii. Mimo że twierdzenie to jest poprawne, zastrzeżenie Zermelo staje się niesłuszne z prostego powodu, a mianowicie czas potrzebny na powrót do stanu początkowego nawet w przypadku niewielkiego pudełka z molekułami gazu jest o wiele

większy od wieku wszechświata. Gdy założymy, że wszechświat przyjął na początku stan niskiej entropii (za czym przemawia doświadczenie), to w dzisiejszym wszechświecie nigdy nie zobaczymy, jak molekuły gazu same zbierają się w jednej części pudełka. Skąd jednak pochodzi ten nieprawdopodobny stan początkowy wszechświata? W tym miejscu pojawia się już kosmologia, a później zobaczymy, że ten problem (i tym samym głębsze uzasadnienie drugiej zasady) może być zrozumiany jedynie w ramach kosmologii.

Także Boltzmann zastanawiał się nad możliwym kosmologicznym pochodzeniem drugiej zasady. Za jego czasów wyobrażano sobie wyłącznie statyczny wszechświat, istniejący wiecznie. Z powodu tego nieskończenie długiego wieku powinno istnieć pewne, choć niezwykle małe, prawdopodobieństwo gigantycznych wahań entropii, jej potężnego spadku i ponownego wzrostu. Ludzkość, jak uważał Boltzmann, istnieje w fazie ponownego wzrostu entropii, czyli w fazie, w której obowiązuje druga zasada termodynamiki. Świat zmierza do swojego stanu równowagi (stanu śmierci cieplnej), który opuści ponownie dopiero po niemierzalnie długim czasie w wyniku wystąpienia ogromnych przypadkowych wahań entropii.

Obecnie wiemy już, że wszechświat powstał w skończonym czasie (zob. część 4). Niezależnie od tego Carl Friedrich von Weizsäcker już w latach trzydziestych XX wieku odkrył słaby punkt w argumentacji Boltzmanna. Zakładając mianowicie, że obserwowany wzrost entropii opiera się wyłącznie na potężnych kosmologicznych wahaniach, wynika z tego, że chwilowe wahania, w ramach których obserwowany świat wraz z ludźmi i ich wspomnieniami powstaje jakby z nicości, są wprawdzie niewyobrażalnie nieprawdopodobne, ale zawsze o wiele bardziej prawdopodobne niż wahania przyjęte przez Boltzmanna. Powód tego polega na tym, że entropia u Boltzmanna musiałaby spontanicznie spaść dużo bardziej, aby stworzyć niezbędny stan wyjściowy, z którego mógłby powstać dzisiejszy stan wszechświata. Byłoby zatem łatwiej bezpośrednio doprowadzić do powstania dzisiejszego stanu. Naturalnie argumentu

Weizsäckera nie należy rozumieć jako potwierdzenia, że powstaliśmy wskutek nieprawdopodobnych wahań entropii, lecz jedynie jako argument przeciwko kosmicznemu scenariuszowi Boltzmanna. W części 6 przekonamy się, co na ten temat mówią teorie kosmologiczne.

Entropia i informacja

Mikroskopowa interpretacja entropii jest w dużej mierze związana z prawdopodobieństwem stanu. Powyższy przykład podzielonego na dwie części pudełka dodatkowo przybliża bliskie pokrewieństwo entropii z informacją. W tamtym przypadku entropia została ostatecznie powiązana z liczbą pytań typu tak lub nie (bity), które są niezbędne do ustalenia określonego stanu molekuł gazu. Im więcej informacji posiadamy, tym mniej pytań musimy zadać i dlatego o tyle mniejsza jest entropia. Amerykański matematyk Claude Shannon (1916–2001) na początku ery komputerowej sformułował pojęcie informacji, które jest ściśle związane z entropią. Entropia to brakująca informacja. W związku z tym ważna staje się zwłaszcza rola istot inteligentnych. Czy ingerencja takich istot w układy termodynamiczne przez użycie informacji może unieważnić drugą zasadę?

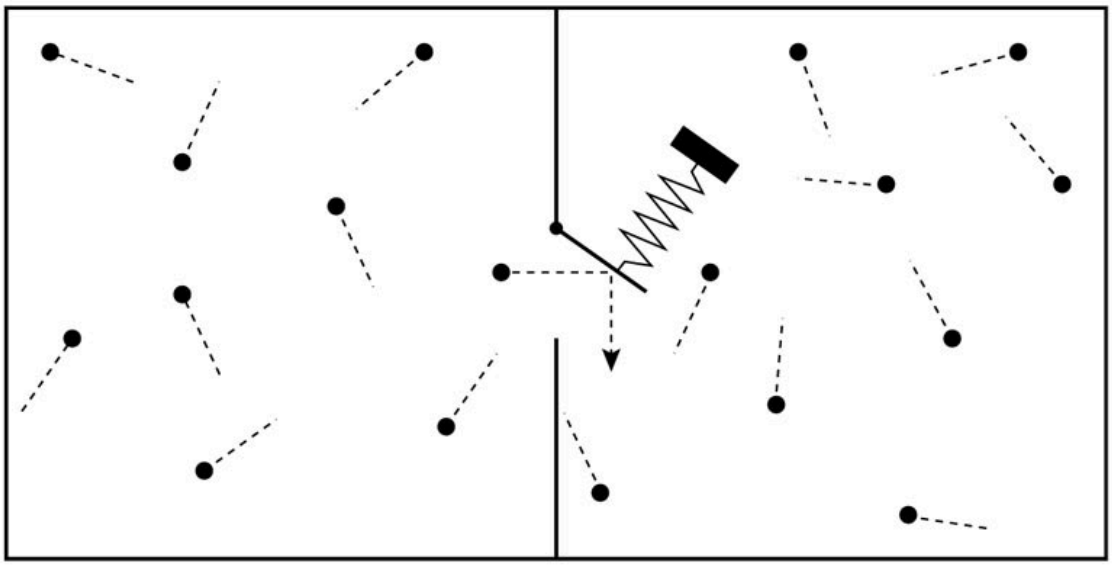
To pytanie ma długą historię. W 1871 roku zajmował się nim szkocki fizyk James Clerk Maxwell, z którym spotkaliśmy się już w części 1. Wyobraził on sobie mikroskopijnie małą istotę, która byłaby w stanie obserwować położenie i prędkość poszczególnych molekuł. Istota ta weszła do historii fizyki pod nazwą „demon Maxwella". Załóżmy ponownie, że pudełko z molekułami gazu jest podzielone ścianką na dwie równe części. W ściance znajduje się niewielki zamykany otwór, przez który poszczególne molekuły mogą przemieszczać się z jednej części do drugiej. Zadanie demona polega na otwieraniu otworu tylko wtedy, gdy szybka molekuła przemieszcza się z prawej strony na lewą, a wolna – z lewej na prawą. Demon przy minimalnym nakładzie pracy mógłby dzięki temu wytworzyć różnicę temperatury między połówkami pudełka i wskutek tego naruszyć drugą zasadę. Oznacza to również, że demon

mógłby na przykład napędzać maszynę parową bez paliwa. Ogromne problemy energetyczne ludzkości byłyby praktycznie rozwiązane. Ostatecznie w oceanach i atmosferze nagromadzona jest ogromna ilość energii termicznej, którą moglibyśmy przetworzyć w pracę, obchodząc drugą zasadę.

Prostszą wersję zaproponował Marian Smoluchowski (1872–1917). W 1912 roku w Lembergu (dzisiejszy Lwów na Ukrainie) wymyślił on mechanizm, w którym wystarczyłoby automatyczne urządzenie – żaden widzący i myślący demon nie byłby potrzebny. Zamiast różnic temperatury konieczne byłoby wytworzenie różnicy ciśnienia bez żadnego nakładu pracy. Można byłoby wtedy bez straty energii obsługiwać na przykład młot pneumatyczny. Pomysł zakłada, że otwór w ściance działowej może być zamykany lub otwierany przez mechanizm zębatkowo-zapadkowy (porównaj z rys. 3.2). Zapadka może być poruszana tylko w jednym kierunku, na przykład w prawo. Molekuły, które przemieszczają się z lewej strony i natrafiają na urządzenie, otwierają zapadkę dzięki ciśnieniu powstałemu podczas zderzenia i pędzą w prawą stronę. Z kolei molekuły przemieszczające się z prawej strony zamykają zapadkę i zostają zamknięte w swojej połowie. Z biegiem czasu w prawej połowie powinno być wyraźnie więcej molekuł, wskutek czego powstanie różnica ciśnienia. W pieczołowicie przygotowanym eksperymencie Smoluchowski odkrył, dlaczego taki pomysł nie może się sprawdzić. Zapadka nagrzewa się podczas zderzeń i po upływie krótkiego czasu przestaje działać w zamierzony sposób. Smoluchowski wyjaśniał to tym, że zawór ma swoje własne prawdopodobieństwo ruchów przypadkowych, a siła jego sprężyny jest albo tak duża, że zawór prawie się nie otwiera, albo tak mała, że cały czas się chwieje i dlatego jest nieskuteczny. W przykładzie Maxwella oznacza to, że demonowi w pewnej chwili zaczynają trząść się ręce i nie może już wypełniać swojego zadania.

Jaki stopień ogólności posiadają te kontrargumenty? Czy istnieje jakaś głębsza przyczyna, dla której istoty inteligentne bądź przemyślane urządzenia automatyczne (które przecież są skonstruowane przez istoty

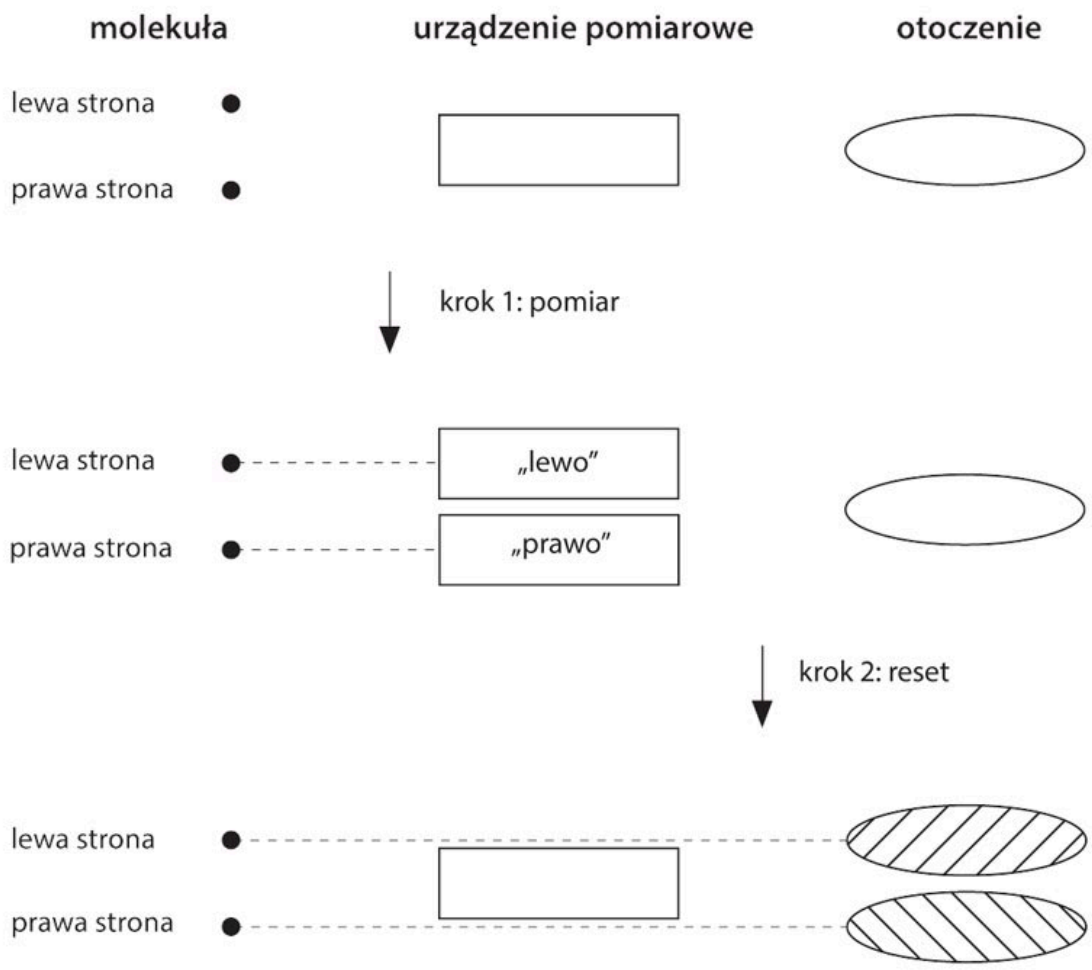
inteligentne) nie mogą naruszać drugiej zasady? W 1929 roku węgierski fizyk Leó Szilárd (1898–1964) stwierdził w ważnym eksperymencie, że podczas obserwacji lub pomiaru, mających na celu zdobycie informacji, musi powstać określona entropia, która w zasadzie wyklucza istnienie *perpetuum mobile* drugiego rodzaju. Należy zatem uwzględnić obserwatora, ponieważ powinien on przejąć (lub przekazać) przynajmniej tę entropię, którą oddaje układ. Gdzie i kiedy dochodzi do powstania entropii? Przez długi czas myślano, że wytworzona entropia siłą rzeczy wynika z obserwacji położenia cząstki. W rzeczywistości przyczyna tkwi gdzie indziej, co w 1973 roku stwierdził pracujący w firmie IBM amerykański fizyk Charles H. Bennett. Aby móc powrócić do stanu wyjściowego i funkcjonować jako część procesu okresowego, demon musi wyczyścić swoją pamięć, co w maszynie odpowiada wyczyszczeniu danych lub resetowi. Wyczyszczenie danych jest nieodwracalne i prowadzi do powstania entropii, która jest oddawana otoczeniu. Notebook nagrzewa się podczas czyszczenia danych (nawet jeśli dzieje się tak w niewielkim stopniu). Pomiar położenia może natomiast przebiegać w sposób odwracalny.



3.2. Eksperyment myślowy Smoluchowskiego

Procesy pomiaru i resetu dla prostej sytuacji są przedstawione na rysunku 3.3. Celem pomiaru jest zdobycie informacji, w której połówce pudełka znajduje się molekuła. Początkowo ani urządzenie pomiarowe, ani otoczenie nic nie wiedzą o położeniu molekuły. W drugim kroku zmianie uległo tylko urządzenie pomiarowe. W zależności od położenia molekuły melduje ono wynik „lewo" lub „prawo". W trzecim kroku urządzenie pomiarowe zostaje ponownie sprowadzone do punktu wyjściowego. Dodatkowo informacja o położeniu molekuły musi wydostać się do otoczenia, co zwiększa jego entropię (więcej możliwych mikrostanów). To, że musi istnieć taki wpływ na otoczenie, wynika z deterministycznej natury dynamiki. Różne stany początkowe nie mogą prowadzić do takiego samego stanu końcowego, ponieważ stan późniejszy jednoznacznie wynika ze stanu wcześniejszego. Informacja ponownie wyczyszczona w urządzeniu pomiarowym musi istnieć dalej w „reszcie świata" (otoczeniu). Nie może być ona jednak wykorzystana w praktyce, z tego względu jest nazywana *nieistotną*.

Entropia ma wiele wspólnego z brakiem wiedzy odnośnie do mikrostanu. Czy z tego powodu pojęcie to jest całkiem subiektywne? Właściwie nie. Definicja entropii wymaga oczywiście wiedzy co do tego, która informacja jest istotna, a która nieistotna. W analogii atramentowej Gibbsa istotne jest to, czy przyglądamy się uważnie i dostrzegamy mikroskopowe nitki atramentu, czy też widzimy jedynie niebieskie zabarwienie wody. W drugim przypadku entropia jest zapewne większa, bo dla tego stanu istnieje znacznie więcej nieznanych mikrostanów, które mogą zrealizować makrostan równomiernego zabarwienia. Wszyscy obserwatorzy mogą jednak wspólnie ustalić definicję istotności i znajdą wtedy dzięki dobrze zdefiniowanym wzorom matematycznym tę samą entropię.



3.3. W celu umożliwienia obiegu informacja o położeniu molekuły musi zostać przekazana do otoczenia (reset), co prowadzi do wzrostu entropii

Wszystkie domniemane paradoksy w rodzaju demona Maxwella zostają obalone, gdy weźmiemy pod uwagę, że informacja musi być opisana zgodnie z zasadami fizyki. Nośniki informacji są zatem układami fizycznymi, które podlegają prawom fizyki. W takim wypadku nie dochodzi do naruszenia drugiej zasady termodynamiki. Informacja nigdy nie może istnieć poza światem. Ludzie jako (względnie nieprawdopodobne) układy o niskiej entropii są w stanie stworzyć w laboratorium nieprawdopodobny stan, na przykład dla gazu, poprzez przejęcie lub odprowadzenie do otoczenia części entropii. Pojawia się

zatem pytanie, dlaczego w ogóle istnieją istoty, które potrafią takie rzeczy, i dlaczego świat od początku nie znajduje się w stanie równowagi termicznej. Pytanie to prowadzi z kolei ponownie do pytania o powstanie i rozwój świata jako całości, a więc do kosmologii, w której łączą się wszystkie ścieżki.

Entropia i teoria kwantowa

W poprzednich dwóch rozdziałach zdefiniowaliśmy entropię poprzez przeliczenie mikroskopowych stopni swobody. W fizyce klasycznej chodzi tu o cząstki (atomy, molekuły), które wykazują własną tożsamość i dlatego można je przeliczyć. W części 2 dowiedzieliśmy się, że na fundamentalnym poziomie nie istnieją żadne cząstki w klasycznym rozumieniu tego słowa. Obowiązywać zaczęła teoria kwantowa, której istotą są funkcje falowe i amplitudy prawdopodobieństwa. Co zatem można przeliczać?

W kwantowej teorii pola pod pojęciem „cząstek", które nie oddziałują ze sobą, rozumie się pola kwantowe, składające się z poszczególnych niezależnych modów drgań o długości fali w przedziale od zera do nieskończoności. Na przykład w pudełku z odblaskowymi ściankami wystąpią tylko takie drgania, które przy ściankach posiadają węzeł (podobną sytuację omawialiśmy w związku z efektem Casimira). Można więc oznaczyć drgania w pudełku poprzez liczbę ich węzłów wraz z węzłami przy ściankach, czyli przez liczby 0, 1, 2 i tak dalej. Każdy z pojedynczych modów drgań drga z własną częstotliwością. W fizyce klasycznej drgania te mogą mieć dowolną amplitudę. W fizyce kwantowej istnieją tylko funkcje falowe, których argumentem są amplitudy. Z funkcji falowych oblicza się następnie rozkład prawdopodobieństwa amplitud.

Z teorii kwantowej wynika następnie, że poszczególne mody drgań mogą przyjmować tylko dyskretne wartości energii, a mianowicie (poza wkładem energii próżni) tylko wielokrotność 0, 1, 2, ... podstawowego kwantu energii. Mimo że właściwie żadne inne cząstki już nie istnieją, kwantyzacja energii symuluje liczbę cząstek, które można przeliczyć.

Jednak istnieje pewna istotna różnica. Określone obsadzenie wszystkich modów kwantami energii uznaje się za tylko jeden stan. Gdyby natomiast chodziło o klasyczne cząstki, mogłyby one jeszcze dodatkowo między sobą permutować bez zmiany makrostanu, co w sumie dałoby oczywiście o wiele więcej stanów. Pomimo rzeczywistego braku cząstek łatwiej jest nadal posługiwać się tym pojęciem przynajmniej w ograniczonym zakresie. Należy jedynie uwzględnić, że te „kwantowe cząstki" są identyczne, czyli nie trzeba dodatkowo zliczać permutacji. Prowadzi to do omawianych już w części 2 bozonów i fermionów – cząstek o spinie całkowitym bądź połówkowym. Cząstki światła (fotony) są przykładem bozonów, ponieważ ich spin = 1. Ich rozkład w pustej przestrzeni jest dany przez prawo promieniowania Plancka, które zapoczątkowało teorię kwantową. Prawo to można wyprowadzić jedynie przez prawidłowe zastosowanie metody przeliczeniowej w odniesieniu do identycznych fotonów. Podczas obliczania entropii jako liczby mikrostanów, które są zgodne z danym makrostanem, należy również uwzględnić tę „identyczność cząstek". Stosując pojęcie cząstki, trzeba pamiętać, że w rzeczywistości chodzi o pola kwantowe, które mogą ze sobą interferować.

Zupełnie nowym aspektem w teorii kwantowej jest omawiane w części 2 splątanie, które nie posiada odpowiednika w fizyce klasycznej. Z uwagi na to, że obserwatorzy mają lokalną naturę, mogą oni doświadczyć jedynie niewielkiej części nielokalnego świata kwantowego, brakuje im zatem informacji na temat tego, co wyraża się w entropii, która nie znika. Entropia jest więc miarą brakującej informacji o stanie splątanym. Przypatrzmy się sytuacji jak w problemie Einsteina, Podolsky'ego i Rosena (EPR). Dwie „cząstki" umykają z jednego źródła w przeciwnych kierunkach, posiadając tylko wspólny stan (wspólną funkcję falową). To wzorcowy przykład splątania. Cząstki te mogą się od siebie oddalać na dowolną odległość. Obserwatorowi, który dostrzega tylko jedną cząstkę, brakuje informacji o splątaniu z innym stanem, dlatego obserwowana cząstka wykazuje dla niego entropię większą od zera.

Obiekty makroskopowe są zazwyczaj splątane ze stopniami swobody swojego otoczenia. Prowadzi to z kolei do omawianego w części 2 procesu dekoherencji, czyli powstania klasycznych właściwości w wyniku oddziaływania z otoczeniem. Dekoherencja jest jednym z najważniejszych nieodwracalnych procesów w przyrodzie. W przypadku stanu, który początkowo nie jest splątany, z biegiem czasu powstaje coraz silniejsze splątanie z innymi układami. Wynika ono z oddziaływania z tymi układami. Wskutek rozproszenia przy udziale światła i molekuł powietrza cząsteczka pyłu osiąga stan splątany z tymi stopniami swobody, przez co zwiększa się lokalna entropia cząsteczki pyłu.

Całkowity układ, który obejmuje wszystkie stopnie swobody, jest tak jak dotychczas w pełni scharakteryzowany przez stan kwantowy i dlatego posiada znikającą entropię. Dopiero dekoherencja podukładów prowadzi do zespołu statystycznego możliwych zdarzeń i związanej z nimi entropii.

W ten sposób zdefiniowana entropia kwantowa również podlega drugiej zasadzie. Z czasem zwiększa się ona albo pozostaje taka sama, nigdy jednak się nie zmniejsza. Ta kwantowa strzałka czasu jest ściśle związana z procesem dekoherencji. O tym, jak bardzo jest ona związana z innymi strzałkami czasu i z wczesnym wszechświatem, przekonamy się w ostatniej części niniejszej książki. Najpierw omówimy jeszcze fundamentalny związek pomiędzy entropią a czarnymi dziurami.

Entropia czarnych dziur

Czarne dziury zachowują się jak układy termodynamiczne. Nie wynikało to z perspektywy ogólnej teorii względności Einsteina. Według niej czarne dziury są masywnymi obiektami z ograniczonym przez horyzont zdarzeń obszarem czasoprzestrzeni, z którego nic, nawet światło, nie może uciec. Takie obiekty mogą powstać jako wynik kolapsu grawitacyjnego gwiazd o dużej masie. W części 2 zobaczyliśmy jednak, że czarne dziury rzeczywiście emitują promieniowanie, jeżeli zostanie uwzględniona teoria kwantowa. Chodzi o słynne promieniowanie Hawkinga, w którym temperatura jest proporcjonalna do kwantu działania Plancka. Zgodnie

z termodynamiką klasyczną, z temperaturą powinna być także związana entropia.

Czy czarna dziura posiada entropię? Na ten temat spekulowano jeszcze przed pracami Hawkinga o promieniowaniu czarnych dziur. Sam Hawking w 1971 roku udowodnił przy ogólnych założeniach, że powierzchnia czarnej dziury (przy czym jest to zawsze powierzchnia horyzontu zdarzeń) nigdy nie może się zmniejszyć. Słychać tu już jakby echo drugiej zasady termodynamiki. Występująca tam entropia również nigdy nie może się zmniejszyć. Z drugiej strony te dwie sytuacje zdają się pochodzić z zupełnie różnych światów. Entropia łączy się w znacznym stopniu z prawdopodobieństwem stanu, podczas gdy twierdzenie Hawkinga dotyczy wielkości całkowicie geometrycznej. Z tego względu związek horyzontu zdarzeń z entropią początkowo nie był traktowany poważnie. Jedynym wyjątkiem był urodzony w 1947 roku fizyk Jacob Bekenstein, wtedy student Johna Wheelera w Princeton. Gdyby czarna dziura nie posiadała entropii, jak przekonywał Wheeler, w prosty sposób zostałaby naruszona druga zasada termodynamiki. Układ wykazujący entropię, jak na przykład filiżankę z gorącą herbatą, wystarczyłoby wrzucić do dziury. Z uwagi na to, że czarna dziura jest oznaczona przez trzy liczby (masę, moment pędu, ładunek), entropia filiżanki z herbatą zanikłaby całkowicie. Całkowita entropia wszechświata zmniejszyła się, a druga zasada została naruszona. Bekenstein potraktował wypowiedź Wheelera niezmiernie poważnie i zaczął wyobrażać sobie bardzo wymyślne eksperymenty myślowe, które miały wykazać, że czarna dziura rzeczywiście posiada entropię. W jednym z nich pudełko wypełnione promieniowaniem zostaje opuszczone na linie ze statku kosmicznego do czarnej dziury. Na horyzoncie zdarzeń pudełko zostaje otwarte, aby promieniowanie mogło przepaść w dziurze na zawsze, po czym pudełko zostaje wciągnięte z powrotem do statku kosmicznego. Po zbadaniu tego procesu Bekenstein wywnioskował, że druga zasada nie zostanie naruszona, jeżeli czarnej dziurze będzie przypisana entropia proporcjonalna do jej powierzchni. Powierzchnia bowiem zwiększa się, gdy materia wpada do dziury (ponieważ wielkość dziury jest proporcjonalna do jej masy). Jak

udowodnił Bekenstein, zwiększenie powierzchni kompensuje co najmniej zmniejszenie entropii w pozostałym wszechświecie, które dochodzi do skutku w wyniku zniknięcia materii w dziurze.

Istnieje zatem związek nie tylko między entropią i powierzchnią horyzontu, w rzeczywistości istnieje także związek ze wszystkimi zasadami termodynamiki. Pierwsza zasada wyraża zachowanie energii. Czarna dziura podlega temu samemu prawu, gdy energia zostanie tylko zastąpiona przez masę dziury pomnożoną przez c^2, co jest logiczne ze względu na równoważność energii i masy zgodnie z teorią Einsteina. Podobnie ma się sprawa z trzecią zasadą. Mimo to na początku lat siedemdziesiątych XX wieku takie związki postrzegano jako całkowicie formalne, ponieważ zgodnie z ogólną teorią względności było nie do pomyślenia, że czarna dziura może emitować promieniowanie. Hawking prezentował wówczas taki sam pogląd. W dążeniu do tego, aby własną pracą obalić wyobrażenie Bekensteina, stwierdził z zaskoczeniem, że czarne dziury rzeczywiście emitują promieniowanie, jeżeli tylko teoria kwantowa będzie stale uwzględniana. Bekenstein miał zatem rację, czarna dziura posiada entropię. Na podstawie swoich wyników Hawking natychmiast określił brakujący współczynnik proporcjonalności, który łączy entropię z powierzchnią. Od tej pory entropię czarnej dziury nazywa się entropią Bekensteina-Hawkinga. W wielu coraz bardziej wyszukanych eksperymentach myślowych stwierdzono, że druga zasada w swojej ogólnej formie pozostaje zachowana. Suma zwykłych entropii i entropii czarnych dziur we wszechświecie nigdy nie maleje.

Jak zmienia się entropia, gdy gwiazda zapada się w czarną dziurę? Przypatrzmy się gwieździe o masie odpowiadającej trzykrotnej masie Słońca. Jej entropia jest poza jedną wartością równa liczbie cząstek, z których się składa. Okazuje się, że czarna dziura, powstająca z tej gwiazdy w wyniku kolapsu, wykazuje entropię, która jest o 20 rzędów wielkości większa. Odpowiada to mnożeniu przez liczbę, która składa się z 1 i 20 zer. W wyniku kolapsu grawitacyjnego dochodzi zatem do ogromnej utraty informacji. Nie jest to zaskakujące, gdyż czarna dziura

jest całkowicie opisana przez trzy wielkości: masę, moment pędu i ładunek – inne informacje nie są dostępne. Dopóki czarna dziura istnieje wraz ze swoim horyzontem zdarzeń, można spokojnie założyć, że informacja jest nadal dostępna w jej wnętrzu, lecz nie można jej zobaczyć z zewnątrz. Czarna dziura emituje jednak promieniowanie Hawkinga, traci przy tym na masie i staje się mniejsza. Co się stanie z informacją, gdy czarna dziura będzie się stale zmniejszać i ostatecznie ulegnie anihilacji? Czy pozornie utracona informacja wydostanie się wtedy na zewnątrz? Według obliczeń Hawkinga promieniowanie czarnej dziury jest oznaczone wyłącznie przez swoją temperaturę. Jeżeli pod koniec procesu parowania pozostałoby jedynie promieniowanie termiczne, informacja o obiektach, które wpadły do dziury, zostałaby ostatecznie utracona. Niezależnie od tego, czy jest to filiżanka, notebook czy encyklopedia, na końcu pozostałoby jedynie promieniowanie. Obliczenia Hawkinga przestają natomiast obowiązywać, gdy czarna dziura jest bardzo mała. W związku z tym pytanie, co stanie się potem, pozostaje nadal otwarte. Dopiero w ramach teorii, która trwale połączy teorię kwantową i teorię względności, będzie można formułować na ten temat hipotezy oraz próbować je udowodnić, dlatego wrócimy jeszcze do tego pytania w części 5. Tam też zastanowimy się, czy można znaleźć mikroskopowe uzasadnienie entropii czarnych dziur zgodnie z założeniami Boltzmanna. Najpierw jednak stworzymy obraz powstania i rozwoju wszechświata zgodnie z ogólną teorią względności i fizyki cząstek elementarnych.

4. Kosmologia

Świat jawi się jako sfera, której centrum
jest wszędzie, a obwód nigdzie.

Mikołaj z Kuzy
(wg przekładu Ireneusza Kani, 1997)

Zasada kosmologiczna

8 kwietnia 1610 roku Johannes Kepler otrzymał *Gwiezdnego posłańca* Galileusza. Zafascynowany opisanymi tam obserwacjami z lunetą, natychmiast zabrał się do sformułowania odpowiedzi. Już 19 kwietnia mógł odesłać Galileuszowi swoją *Dissertatio cum Nuncio Sidero* (Rozmowa z Gwiezdnym Posłańcem). W pracy tej odnosi się on do wielu nowo odkrytych gwiazd:

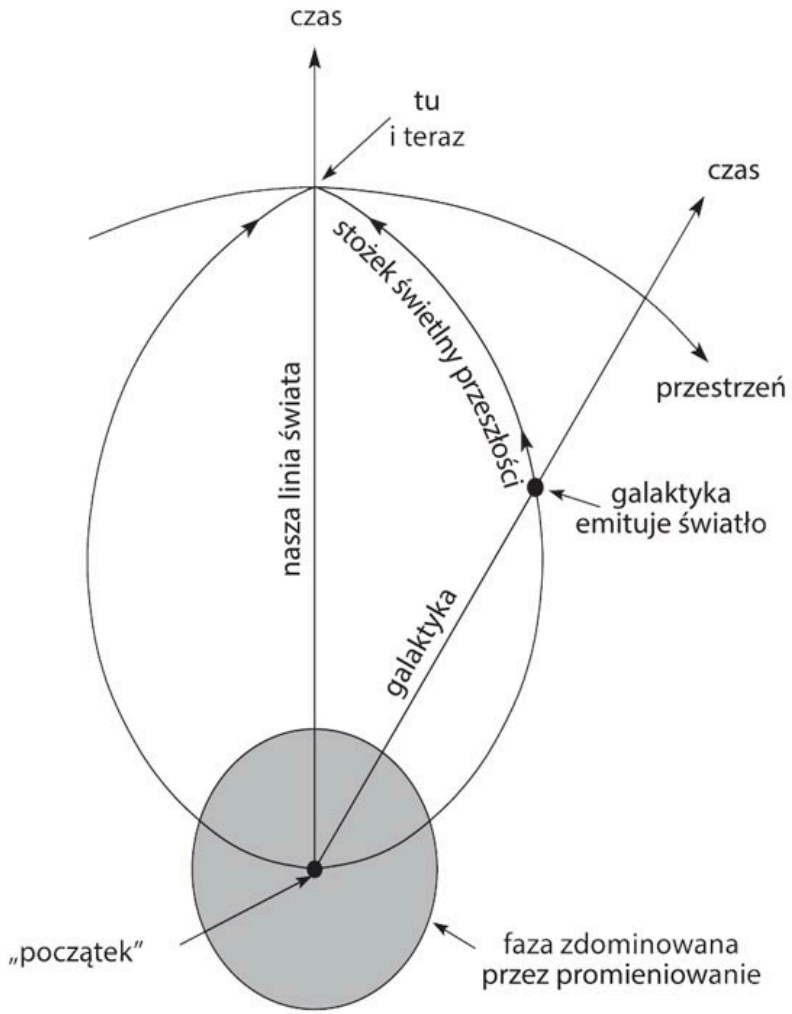
> Jeżeli to prawda, i jeżeli owe słońca są tej samej natury co nasze, dlaczego te wszystkie słońca nie prześcigną blaskiem naszego Słońca?

Jak już wspomnieliśmy w części 1, Kepler odniósł się tutaj do potencjalnej nieskończoności wszechświata. Jeżeli za pomocą pierwszej i najprostszej lunety można już było zobaczyć wiele nowych gwiazd, czy z każdym kolejnym etapem technicznego rozwoju nie będzie widać ich jeszcze więcej? Czy oznacza to, że wszechświat jest nieskończenie wielki?

Założenie to prowadzi natychmiast do problemu poruszonego przez Keplera. Nieskończenie wielki wszechświat, równomiernie wypełniony gwiazdami, powinien skutkować oślepiająco jasnym „nocnym niebem". Spoglądając w głąb lasu, widać wszędzie drzewa. Analogicznie w nieskończonym wszechświecie gwiazdy powinny być widoczne wszędzie. Taki pojęciowy problem jest znany jako *paradoks Olbersa*,

a owa nazwa pochodzi od nazwiska bremeńskiego lekarza i astronoma Heinricha Wilhelma Olbersa (1758–1840), który opisał go niezależnie od Keplera. Nierozwiązany w ramach kosmologii Newtona, paradoks ten znajduje swoje naturalne wyjaśnienie w ogólnej teorii względności. Gwiazdy zaczęły świecić przed pewnym skończonym okresem czasu, a ich światło ze względu na ekspansję wszechświata ulega osłabieniu (przesunięcie ku czerwieni). Dowody na to przedstawimy poniżej.

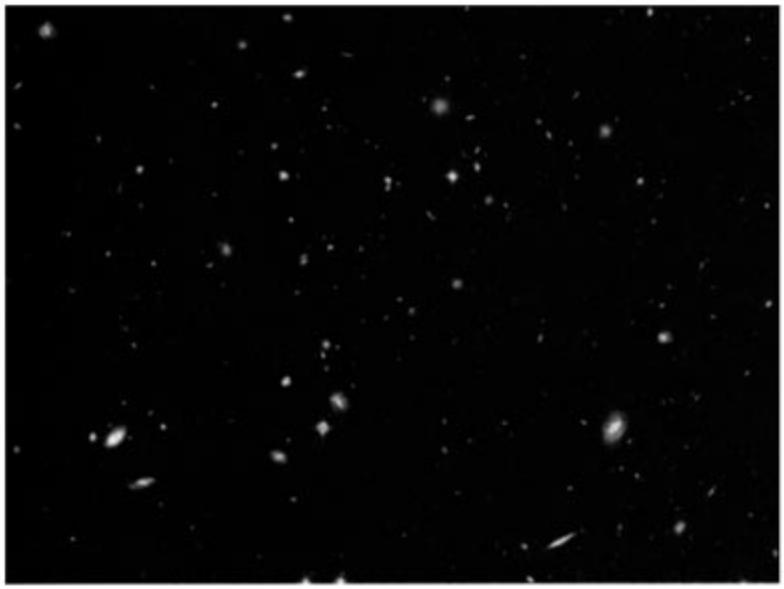
Patrząc dziś w kosmos, zderzamy się z pewnego rodzaju granicą widzialności. Musimy sobie przy tym uzmysłowić, że światło przemierza wszechświat ze skończoną prędkością, a my widzimy odległy obiekt nie w stanie obecnym, lecz w stanie wcześniejszym, w którym światło zostało wysłane. Patrzymy zatem nie tylko na przestrzenne, ale i czasowe odległości. Pokazano to na rysunku 4.1. Widać na nim stożek świetlny przeszłości sięgający od tu i teraz w przeszłość (z uwagi na to, że przedstawiono tylko jeden wymiar przestrzeni, „stożek świetlny przeszłości" jest tylko krzywą, która pozostaje zakrzywiona, ponieważ wszechświat się rozszerza).



4.1. Stożek świetlny przeszłości, z którego pochodzi praktycznie każda nasza informacja z przeszłości. Przedstawiono także naszą linię świata i linię obserwowanej przez nas galaktyki.

Prawie każda informacja o wszechświecie dociera do nas za pośrednictwem tego stożka. Niewiele wyjątków dotyczy badań skamieniałości i skał meteorytów na Ziemi, jak również występowania cząstek promieniowania kosmicznego. Bezpośrednio możemy więc zaobserwować jedynie niewielki ułamek czasoprzestrzeni! Na podstawie obserwacji obcych galaktyk tylko wtedy możemy nauczyć się czegoś o stanie przeszłym naszej własnej Drogi Mlecznej, gdy założymy, że przestrzeń wszechświata jest dość jednorodna. Okazuje się wówczas, że

nasza Droga Mleczna w odległej przeszłości była mniej więcej w takim stanie jak podobna galaktyka, której światło wysłane w równie odległej przeszłości dociera do nas dziś.



4.2. *Ultra Deep Field* w ujęciu teleskopu Hubble'a

Rysunek 4.2 to spojrzenie w przeszłość wszechświata. Przedstawia on *Ultra Deep Field*, uzyskane za pomocą kosmicznego teleskopu Hubble'a w ujęciu trwającym trzy miesiące. To daleko sięgające zdjęcie ukazuje niewielki wycinek nieba. Potrzebowalibyśmy pięćdziesięciu takich ujęć, aby pokryć powierzchnię Księżyca. Z wyjątkiem niewielu gwiazd, które można rozpoznać po ich migotaniu, widać tylko galaktyki. Spoglądamy w przeszłość, do chwili powstania pierwszych galaktyk we wczesnym wszechświecie. W tym sensie osiągnęliśmy granice widzialnego wszechświata, przynajmniej jeśli chodzi o światło gwiazd.

Wspominaliśmy już o przestrzennej jednorodności wszechświata. Czy jest to rozsądne? Ostatecznie wszędzie widzimy różnorodne struktury, z których nie można wywnioskować, że wszechświat wygląda wszędzie tak samo. Jeśli mówimy o jednorodnym wszechświecie, mamy na myśli to, że w bardzo dużych skalach, biorąc pod uwagę średnie wielkości wielu

odległości międzygalaktycznych, jest on mniej więcej jednorodny. Czy istnieją ku temu przesłanki empiryczne?

Od 1965 roku wiemy, że Ziemia jest otoczona słabym promieniowaniem termicznym, którego temperatura około 2,7 kelwina leży tuż powyżej zera absolutnego. Promieniowanie to jest w wysokim stopniu izotropowe, to znaczy, że we wszystkich kierunkach wygląda praktycznie tak samo. Odchylenia od izotropii wynoszą jedynie jedną stutysięczną (co będzie miało znaczenie dopiero w późniejszej części tekstu). Promieniowanie nie ma zatem swojego początku w dyskretnych źródłach i już wcześniej istniała pewność co do tego, że pochodzi ono z kosmosu. Nazwa „kosmiczne promieniowanie tła" również wskazuje na to pochodzenie. Jest to pozostałość po bardzo gorącym promieniowaniu, które wraz z rozszerzaniem się wszechświata schłodziło się do zaledwie 2,7 kelwina. Istnienie tego promieniowania jest dowodem na izotropię wszechświata. Fakt ten znajduje dodatkowe potwierdzenie w rozmieszczeniu słabych radiogalaktyk oraz tła wysokoenergetycznego promieniowania gamma. Niestety ten sam obraz we wszystkich kierunkach nie dowodzi jeszcze, że przestrzeń wszechświata jest jednorodna. Moglibyśmy na przykład znajdować się w centrum niejednorodnego, ale sferycznie symetrycznego wszechświata. Nie posuniemy się jednak naprzód bez dodatkowej zasady lub postulatu. Goethe pisał do kompozytora Zeltera, że największą sztuką w nauczaniu i światowym życiu jest umiejętność zamienienia problemu w postulat, bo wtedy można zrobić krok naprzód. To, dzięki czemu można zrobić krok naprzód w tej kwestii, to tak zwana *zasada kosmologiczna*, nazywana także zasadą kopernikańską. Postulujemy, że wszechświat z każdego miejsca (nie tylko z Ziemi) we wszystkich kierunkach prezentuje ten sam widok. Ziemia nie zajmuje w nim więc specjalnego miejsca. Twierdzenie matematyka Friedricha Schura, udowodnione w 1886 roku, głosi, że z izotropii wokół jednego punktu wynika przestrzenna jednorodność. Zakładamy zatem, że wszechświat jest jednorodny, biorąc pod uwagę wielkie przestrzenie.

Mimo to warto przyjrzeć się strukturom zaobserwowanym we wszechświecie. Stosowaną przy tym zwykle jednostką długości jest parsek, a dokładniej megaparsek (milion parseków), czyli w skrócie Mpc. Jak to często bywa w astronomii, parsek ma swój historyczny rodowód. Został zdefiniowany jako odległość od Słońca, z której promień orbity ziemskiej widać pod kątem jednej sekundy kątowej. Jeden parsek to nieco ponad trzy lata świetlne lub około 31 bilionów kilometrów.

Najmniejszymi tworami w kosmologii są galaktyki. Wzorcowy przykład stanowi nasza własna Droga Mleczna, mgławica spiralna. Chodzi o dysk z wypukłością w środku o masie całkowitej przekraczającej 100 miliardów mas Słońca (100 miliardów to także przybliżona liczba neuronów w ludzkim mózgu). Średnica dysku wynosi około 30 kiloparseków (30 000 parseków), przy czym nasz Układ Słoneczny jest umiejscowiony raczej na skraju, oddalony od środka o niecałe 10 kiloparseków. Droga Mleczna jest otoczona sferoidalnym halo, w którym znajdują się przede wszystkim gromady gwiazd kulistych, należących do najstarszych obiektów naszej galaktyki.

W widzialnej części wszechświata istnieje z kolei około 100 miliardów takich galaktyk, które łączą się w gromady i supergromady. Gromady mają rozmiar około 5 Mpc, a supergromady – 50 Mpc. Ponadto nie ma już żadnych rozpoznawalnych superstruktur. Biorąc pod uwagę średnie wartości odległości mniej więcej od 10 do 100 Mpc, wszechświat wydaje się w tej skali jednorodny. Rozmiar widzialnego wszechświata wynosi około 4000 Mpc. Między supergromadami istnieją puste przestrzenie (*Voids*) o rozmiarze od około 100 do 200 Mpc. Nasza Droga Mleczna należy do gromady zwanej *Grupą Lokalną*, obejmującą ponad 30 galaktyk, jak również do supergromady nazwanej *Supergromadą w Pannie*, której centrum znajduje się w odległości około 20 Mpc od nas. Z uwagi na to, że cała objętość wypełniona galaktykami stanowi mniej niż jedną milionową całkowitej objętości wszechświata, uzasadnione jest w kosmologii (i tylko tu) postrzeganie galaktyk jako „punktów”. W dalszej odległości znajdują się jeszcze kwazary – obiekty emitujące silne promieniowanie,

w przypadku których mamy prawdopodobnie do czynienia z galaktykami we wczesnym stadium ich rozwoju, czyli poprzednikami obecnych galaktyk. Ich silne promieniowanie jest wynikiem oddziaływania z materią, która wpływa do supermasywnej czarnej dziury znajdującej się w centrum kwazara. Wskutek tego energia grawitacyjna tej materii zostaje uwolniona i wyemitowana.

Przez długi czas wierzono, że we wszechświecie istnieje tylko taka materia, którą można dostrzec (poprzez emitowane przez nią promieniowanie elektromagnetyczne). Tak jednak nie jest. Już w latach trzydziestych XX wieku szwajcarsko-amerykański astronom Fritz Zwicky (1898–1974) na podstawie obserwacji dynamiki gromad galaktyk doszedł do wniosku, że musi istnieć bardzo dużo materii, której nie można zaobserwować – „ciemna materia". Ze względu na to, że galaktyki w takich gromadach mają zbyt wielkie prędkości, musi to oznaczać, że w obecności wyłącznie widzialnej materii galaktyki już dawno oddaliłyby się od gromad. Musi więc istnieć dodatkowa materia, która dzięki swojej grawitacji może utrzymać galaktyki na ich orbitach wokół centrum. Podobne wnioski nasuwają się na podstawie badań prędkości gwiazd w poszczególnych galaktykach spiralnych. Niestety o naturze ciemnej materii niewiele jeszcze wiadomo, ale o tym później. Zadaniem astronomów zajmujących się obserwacją jest sporządzenie dokładnej kartografii *cosmic web*, czyli sieci kosmicznej, składającej się z widzialnej i niewidzialnej materii. Astronomowie, w odróżnieniu od ich pływających po morzach poprzedników, nie muszą wsiadać w tym celu do statków kosmicznych, lecz jedynie konstruować coraz lepsze teleskopy.

Odkrycie, że galaktyki to systemy poza naszą Drogą Mleczną, nie jest wcale takie stare. Zyskało popularność dopiero na początku lat dwudziestych XX wieku, po tym jak amerykański astronom Edwin Hubble (1889–1953) na podstawie dowodu dotyczącego gwiazd zmiennych w Mgławicy Andromedy i innych galaktykach mógł ustalić ich odległość. Ten sam Hubble ustalił w 1929 roku jeszcze coś innego, a mianowicie, że odległe galaktyki wykazują w swoich widmach

przesunięcie ku czerwieni, które jest proporcjonalne do ich odległości. Takie przesunięcia ku czerwieni można było interpretować wyłącznie jako efekt Dopplera (porównaj z częścią 1), co oznacza, że galaktyki te oddalają się od nas. Czy zatem znajdujemy się jednak w centrum wszechświata? W średniowieczu odkrycie to zostałoby wytłumaczone zapewne w taki właśnie sposób. W ramach ogólnej teorii względności związek między przesunięciem ku czerwieni a odległością interpretuje się jako samą ekspansję przestrzeni. Galaktyki jak małe szpilki umocowane są w stałych punktach przestrzeni, a przestrzeń między galaktykami rozszerza się. Popularną analogią jest nadmuchany balon, na którym galaktyki zostały zaznaczone jako punkty. Patrząc z perspektywy każdego punktu, podczas nadmuchiwania balonu wszystkie punkty oddalają się, i to tym szybciej, im bardziej są oddalone. Wszechświat rozszerza się.

Stałą proporcjonalności między przesunięciem ku czerwieni (lub pozorną prędkością) a odległością nazywa się na cześć Hubble'a – stałą Hubble'a. Chodzi tu o jedną z najważniejszych wielkości w kosmologii, wszak jest miarą prędkości ekspansji wszechświata. Jednym z głównych zadań kosmologii obserwacyjnej jest ustalenie tej wielkości. O istnieniu związku między przesunięciem ku czerwieni a odległością, jeszcze zanim uczynił to Hubble, mówił niemiecki astronom Carl Wilhelm Wirtz (1876–1939). W 1924 roku napisał on, że nie ma już żadnych wątpliwości co do tego, iż dodatni ruch radialny mgławicy spiralnej znacznie wzrasta wraz ze zwiększającą się odległością. Wirtz w przeciwieństwie do Hubble'a nie dysponował jednak precyzyjnymi teleskopami, a więc nie mógł obliczyć tego związku.

Jeśli wszechświat się rozszerza, to w przeszłości musiał być mniejszy i powstać przed pewnym skończonym okresem czasu. Z uwagi na to, że jest wypełniony kosmicznym promieniowaniem tła, które ulega ochładzaniu wraz z jego rozszerzaniem, w swoich początkach musiał być gorący. Z tego powodu mówi się o „gorącym wielkim wybuchu", od którego wszechświat wziął swój początek. Takie pojęcie wprowadził w 1949 roku brytyjski astronom Fred Hoyle (1915–2001), traktował je

jednak jako obelżywe określenie nielubianej przez siebie teorii ekspansji. Przesunięcie ku czerwieni łącznie ze skończonym wiekiem wszechświata prowadzi do rozwiązania wspomnianego na wstępie paradoksu Olbersa.

Przykład rozszerzającego się balonu może łatwo wprowadzić w błąd. Podczas gdy balon rozszerza się w normalnej trójwymiarowej przestrzeni obserwacji, w przypadku wszechświata to sama przestrzeń się rozszerza. Całkowicie w myśl geometrycznego rozwoju od czasów Gaussa i Riemanna liczy się nie osadzenie w większej przestrzeni, ale wewnętrzna geometria. Niezależnie od tego, czy istnieje jakaś większa przestrzeń (jak na przykład w teoriach wyższych wymiarów), wszystkie wielkości tej teorii i wszystkie obserwacje dotyczą jedynie tej przestrzeni. Nie możemy przecież opuścić wszechświata i popatrzeć na niego z zewnątrz.

Czy ekspansja wszechświata oznacza, że wszystkie skale także ulegają rozszerzeniu, czyli że na przykład przedmioty w naszym pokoju lub atomy w naszym ciele oddalają się od siebie? W żadnym razie. Struktury połączone ze sobą zachowują swoje wzajemne odległości i biorą udział w ekspansji jako jedność. Na przykład galaktyki w naszym sąsiedztwie (członkowie wymienionej Grupy Lokalnej) są między sobą i z Drogą Mleczną tak mocno związane grawitacyjnie, że rozszerzają się wspólnie. Dopiero w większych skalach ekspansja wszechświata prowadzi rzeczywiście do zwiększenia się odległości między gromadami galaktyk.

Rozszerzający się wszechświat, przestrzennie jednorodny i izotropowy, musiał powstać przed pewnym skończonym okresem czasu w wyniku „wielkiego wybuchu”. W pojęciach ogólnej teorii względności stan ten odpowiada osobliwości, czyli stanowi o nieskończenie wielkiej gęstości i krzywiźnie. Z tego powodu długo przypuszczano, że osobliwość zniknie, gdy upadnie założenie jednorodności i izotropii, które wcale nie musi obowiązywać we wczesnym wszechświecie. Wspomniane w części 1 twierdzenia dotyczące osobliwości Penrose’a, Hawkinga i innych uczą, że do wielkiego wybuchu doszło w dość ogólnych warunkach, niezależnie od przyjęcia jakiejkolwiek symetrii. Obserwacje przemawiają za tym, że te ogólne warunki są spełnione. Przeszłość wszechświata jest więc względnie

znana. W przeciwieństwie do tego (jak zwykle) trudniej jest przewidzieć przyszłość. Tym także zajmiemy się później. Najpierw omówimy szczegóły modeli jednorodnych i izotropowych, które wynikają z założeń teorii względności.

Wszechświat Friedmana i Lemaître'a

Godzina narodzin nowoczesnej kosmologii wybiła wraz z pracą Einsteina pod tytułem *Kosmologiczne rozważania nad ogólną teorią względności* (wg przekładu Michała Hellera, 2005), którą opublikował w 1917 roku w sprawozdaniach Królewsko-Pruskiej Akademii Nauk w Berlinie. To tu po raz pierwszy zastosował on swoją teorię do wszechświata jako całości. Jak już wspomnieliśmy w części 1, w tym czasie Einstein należał do zwolenników zasady Macha, według której zawartość materii we wszechświecie powinna ustalać jednoznacznie geometrię przestrzeni, a przede wszystkim nie powinien istnieć pusty wszechświat. Einstein już wtedy zakładał jednorodność i izotropię wszechświata, którą szeroko uzasadniliśmy w ostatnim rozdziale. Potrafił on jednak wyobrazić sobie jedynie statyczny, czyli niezmienny w czasie wszechświat. Jako model przestrzeni wybrał trójwymiarową sferę, czyli odpowiednik dwuwymiarowej powierzchni kuli w zwykłej przestrzeni trójwymiarowej. Podobnie jak powierzchnia kuli przestrzeń ta jest nieograniczona, ale skończona – posiada skończoną objętość. Z uwagi na to, że nie ma krawędzi, nie istnieje również swoboda co do wyboru warunków brzegowych. W związku z tym powinno dojść do implementacji zasady Macha, a sama zawartości materii powinna decydować o geometrii.

Pojawił się jednak jeden problem. Równania Einsteina w formie z 1915 roku nie dopuszczały istnienia statycznego wszechświata. Musiał się on rozszerzać albo zapadać, czego Einstein nie chciał zaakceptować. Stwierdził, że jego równania nie przyjęły jeszcze najbardziej ogólnej formy i można wprowadzić jeden dodatkowy parametr. Parametrem tym jest słynna stała kosmologiczna. W odróżnieniu od zwykłej grawitacji prowadziła ona do odpychania, które jest odczuwalne dopiero przy bardzo

dużych (właśnie kosmicznych) odległościach. W związku z tym stała kosmologiczna może stawić opór przyciągającemu oddziaływaniu materii. Dla określonej wartości przyciąganie i odpychanie utrzymują się na równym poziomie, a wynikający z tego wszechświat jest statyczny. Oto model świata według Einsteina z 1917 roku. Im mniej materii istnieje w tym wszechświecie, tym jest on mniejszy. W granicznym przypadku pustego świata znika jego promień, czyli pusty świat nie mógł istnieć – całkiem w duchu zasady Macha.

Model Einsteina nie utrzymał się długo. Jeszcze w tym samym 1917 roku niderlandzki astronom Willem de Sitter (1872–1934) opublikował własne rozwiązanie równań Einsteina ze stałą kosmologiczną, które poradziło sobie bez materii i tym samym zaprzeczyło zasadzie Macha. Einstein szybko stwierdził, że zasada w postulowanej przez niego formie nie jest słuszna i istnieje wiele różnych rozwiązań równań Einsteina z uwzględnieniem materii, jak i bez niej, i dlatego z niej zrezygnował. Potem nadmienił, że w zasadzie nie powinno się więcej wspominać o zasadzie Macha.

Kilka lat później rosyjski fizyk Aleksander Friedman (1888–1925) opublikował w „Zeitschrift für Physik" dwie znaczące prace, w których przedstawił rozwiązania równań Einsteina ze stałą kosmologiczną i bez niej. Opisują one jednorodny i izotropowy wszechświat, który albo się rozszerza i ponownie zapada, albo rozszerza się na zawsze. Prace Friedmana zostały wydane w 1922 i 1924 roku i nie odwoływały się do żadnych obserwacji. Krótko po tym belgijski kosmolog i ksiądz katolicki Georges Lemaître (1894–1966) sprawdził te rozwiązania i znalazł powiązanie przede wszystkim z zaobserwowanym przesunięciem ku czerwieni galaktyk. Jako pierwszy wyszedł z założenia, że wszechświat rozwinął się w skończonym czasie z bardzo gęstego i gorącego stanu początkowego, dlatego słusznie nazywany jest ojcem teorii wielkiego wybuchu.

Roztrząsane przez Friedmana i Lemaître'a rozwiązania zwane są na ich cześć najczęściej rozwiązaniami lub równaniami Friedmana-

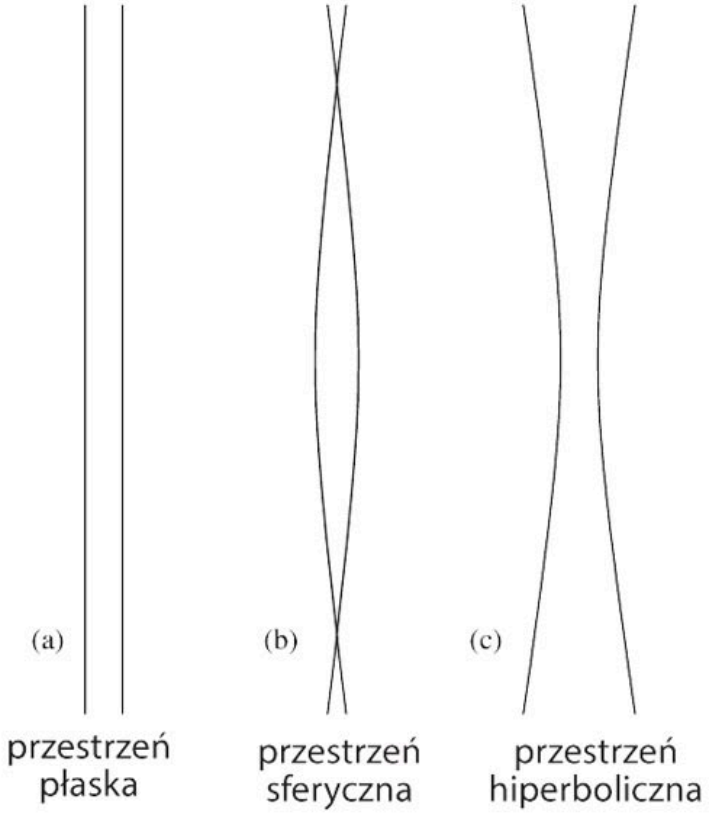
Lemaître'a (czasem po prostu tylko równaniami Friedmana). Opisują one w myśl zasady kosmologicznej przestrzenny wszechświat, który jest jednorodny i izotropowy. Teraz skupimy się na prostszym przypadku, w którym stała kosmologiczna znika. Bardziej ogólny przypadek zostanie szczegółowo omówiony w jednym z kolejnych rozdziałów.

Jednorodne i izotropowe przestrzenie są przestrzeniami o stałej krzywiźnie – w każdym punkcie przestrzeni zależności geometryczne są takie same. Jakie możliwości istnieją w ogóle w takim przypadku? Analizy matematyczne uczą, że przestrzenie o stałej krzywiźnie można podzielić na trzy kategorie: z zerową, dodatnią lub ujemną krzywizną. Charakteryzuje je parametr k, który może przyjąć wartości 0, 1 lub –1. Dla każdego z tych trzech przypadków istnieje nieskończenie wiele możliwości. W razie postulowania kolejnej właściwości dla każdego przypadku będzie istniało dokładnie jedno rozwiązanie. Przestrzenie powinny posiadać pewną właściwość: powinny być jednospójne. Jednospójne przestrzenie charakteryzują się tym, że każdą zamkniętą w nich krzywą można ciągle ściągnąć do punktu. Mówiąc obrazowo, przestrzeń nie powinna posiadać żadnych dziur czy uchwytów, na których krzywa mogłaby się „zawiesić".

Czym są te trzy pozostałe przypadki? W przypadku $k = 0$ uzyskuje się płaską trójwymiarową przestrzeń, w której obowiązują prawa geometrii euklidesowej. Wówczas dla każdej prostej i każdego punktu poza prostą istnieje dokładnie jedna równoległa do niej prosta. W przypadku $k = 1$ jest to przestrzeń, z której korzystał już Einstein w swoim modelu z 1917 roku, czyli trójwymiarowa sfera. Tu nie ma żadnej równoległej, ponieważ wszystkie geodezyjne o tym samym kierunku początkowym zbliżają się do siebie i ostatecznie przecinają. Przypadek ten jest najciekawszy, gdyż stwarza możliwość wyobrażenia sobie skończonego, ale przestrzennie nieograniczonego wszechświata. W przypadku $k = -1$, tak zwanej pseudosfery, geodezyjne o tym samym kierunku początkowym oddalają się od siebie i dlatego istnieje nieskończenie wiele równoległych. Objętość

jest nieskończenie wielka. Na rysunku 4.3 przedstawiono te trzy możliwości.

Jeśli odrzucimy postulat jednospójności, to efektem tego będzie nieskończenie wiele, po części bardzo dziwnych, przestrzeni o stałej krzywiźnie. Właśnie im poświęcony jest kolejny rozdział na temat topologii wszechświata. W tym miejscu ograniczymy się do trzech najprostszych możliwości.



4.3. Równoległe w przestrzeni płaskiej (a), przestrzeni sferycznej (b) i przestrzeni hiperbolicznej (c). W przestrzeni płaskiej dla każdej prostej istnieje jedna równoległa przecinająca punkt leżący poza prostą, w przestrzeni sferycznej nie ma żadnej równoległej, a w hiperbolicznej jest ich nieskończenie wiele.

Klasyfikacja ta dotyczy całkowicie przestrzennego aspektu, czyli stanu wszechświata w stałym czasie, i wynika bezpośrednio z zasady

kosmologicznej. Jeżeli dojdzie do tego jeszcze czas, powstanie pewien obraz w układzie dynamicznym. Z równań Einsteina wynika, że wszechświat rozszerza się lub kurczy. Jednak co dokładnie rozszerza się lub kurczy? Przyjrzyjmy się punktom („galaktykom"), które są przyczepione do stałego punktu w przestrzeni. W jednorodnym i izotropowym wszechświecie rozszerzanie się przestrzeni jest wszędzie równomierne. Może ono być opisane przez jedną wielkość, czyli tak zwany czynnik skali. Wskazuje on, w jaki sposób z biegiem czasu zmienia się odległość między dwiema galaktykami. Nieco potocznie mówi się także o „promieniu wszechświata", lecz takie określenie jest adekwatne najwyżej w przypadku sfery, której objętość jest skończona. W odniesieniu do nieskończonej objętości pojęcie promienia nie ma sensu.

Gdy stała kosmologiczna znika, jak tu przyjęto, każdej z trzech przestrzennych krzywizn odpowiada dokładnie jeden rozwój w czasie. Przypadek sfery ($k = 1$) tworzy wszechświat, który począwszy od wielkiego wybuchu rozszerza się, osiąga maksimum, a następnie ponownie się kurczy i po upływie skończonego odcinka czasu dochodzi do odpowiednika wielkiego wybuchu, czyli „wielkiego krachu". W przypadku pseudosfery ($k = -1$) nieskończenie wielki wszechświat rozszerza się począwszy od wielkiego wybuchu przez cały czas, przy czym czynnik skali dla dużych odcinków czasu wzrasta liniowo wraz z czasem. Przypadek graniczny między sferą a pseudosferą tworzy płaska przestrzeń ($k = 0$). Nieskończenie wielki wszechświat rozszerza się przez cały czas, przy czym znika prędkość dla dużych odcinków czasu. Naturalnie w dwóch ostatnich przypadkach wszechświat mógłby rozwijać się w odwrotnym kierunku upływu czasu, czyli począwszy od nieskończoności mógłby się stale kurczyć, aż ostatecznie zniknąłby w wyniku wielkiego krachu. Równania Einsteina nie wyróżniają żadnego kierunku upływu czasu na tle innych. To zaobserwowana ekspansja wszechświata jest tym, co wyklucza taką ewentualność.

W trzech opisanych powyżej przypadkach pobrzmiewa jeszcze echo kroków Newtona. W mechanice nieba także istnieją trzy przypadki. Gdy

wskutek działania siły ciężkości ciało okrąża jakieś inne ciało, przy naszym założeniu o wiele cięższe, ruch ten może odbywać się po elipsie, paraboli lub hiperboli, w dodatku w zależności od energii ciała. W przypadku energii ujemnej ruch jest kontrolowany, a ciało biegnie periodycznie po elipsie wokół drugiego ciała, czego przykładem jest ruch planet wokół Słońca. W przypadku energii dodatniej ruch nie jest kontrolowany, ciało przybywa z nieskończoności, krąży wokół innego ciała i znika bezpowrotnie we wszechświecie. Krzywą orbity jest hiperbola. Za przykład z naszego Układu Słonecznego mogą posłużyć pewne komety, które odwiedzają nas tylko jeden raz, tak jak wspomniana w części 1 kometa McNaughta. (Inne komety, jak słynna kometa Halleya, biegną po elipsach i dlatego odwiedzają nas w regularnych odstępach czasu). Przypadkiem granicznym między elipsą a hiperbolą jest parabola, gdy energia znika, a prędkość przy dużych odległościach zbliża się do zera. Między kosmologią Newtona a Einsteina istnieją jednak istotne różnice. Podczas gdy u Newtona rozróżnia się energię, u Einsteina ważny jest znak krzywizny, czyli kryterium całkowicie geometryczne. Dodatkowo to nie wszechświat rozszerza się *do* czegoś, ale sama całkowita przestrzeń staje się większa.

Może nieco dziwić, w jaki sposób nieskończenie wielka przestrzeń jest w stanie się rozszerzać. Nawet jeśli nasze wyobrażenie wzbrania się przed pojęciem nieskończoności, opis matematyczny nie ma z tym żadnego problemu. Wyobraźmy sobie nieskończoną przestrzeń wypełnioną równomiernie nieskończenie wieloma galaktykami. Z biegiem czasu odległość między dwiema dowolnymi galaktykami będzie się zwiększać, co w nieskończenie wielkiej przestrzeni jest możliwe. Odpowiada to ekspansji wszechświata. Nawet podczas wielkiego wybuchu przestrzeń jest nieskończenie wielka. Odległość między galaktykami (o ile takie by jeszcze istniały) znika, a ta nieskończenie wielka przestrzeń składa się wtedy z jednej osobliwości. Tylko w przypadku sfery początkowa osobliwość składa się z jednego punktu. Dla nas obserwatorów dostępny jest jednak tylko skończony obszar wszechświata.

Równania Einsteina formułują związek geometrii z materią. W przypadku modeli Friedmana-Lemaître'a prowadzą one do prostego stosunku między gęstością materii a wielkością k, która charakteryzuje trzy powyższe przypadki. Oczywiście w takim modelu rozkład materii również jest jednorodny, ponieważ w przeciwnym razie z równań wynikałaby niejednorodna geometria będąca zaprzeczeniem zasady kosmologicznej. Okazuje się, że istnieje tak zwana „gęstość krytyczna", której odpowiada przypadek $k = 0$, czyli przypadek płaskiej przestrzeni. Gdy gęstość jest mniejsza, mamy do czynienia z przypadkiem $k = -1$, a gdy większa – z przypadkiem $k = 1$. Z obserwacji rozkładu materii we wszechświecie można w zasadzie wyciągać wnioski co do jego geometrii i tego, czy jest on skończenie, czy nieskończenie wielki. W kontekście tym pomocne jest wprowadzenie tak zwanego parametru omega (Ω) wszechświata. Jest on zdefiniowany jako stosunek gęstości do gęstości krytycznej. Gdy parametr ten jest większy niż jeden, to mamy do czynienia z wszechświatem $k = 1$, gdy jest mniejszy niż jeden – z wszechświatem $k = -1$. W przypadku granicznym $\Omega = 1$ przestrzeń wszechświata jest płaska. Związek ten jest zachowany także z biegiem czasu. Parametr k i tym samym zależności geometryczne pozostają stałe. Dalej dowiemy się, co obecnie wiadomo o parametrze Ω.

W ramach modeli Friedmana-Lemaître'a można uzasadnić prawo Hubble'a, czyli związek między odległością a przesunięciem ku czerwieni. W przypadku niewielkich przesunięć ku czerwieni, czyli bliskich galaktyk, widoczny jest liniowy związek między tymi wielkościami, odkryty przez Hubble'a. Podwójna odległość odpowiada podwójnemu przesunięciu ku czerwieni. Do odchyleń dochodzi przy większych przesunięciach ku czerwieni, to jest w przypadku bardziej odległych obiektów. Przyczyną tego jest rozwój wszechświata w czasie między momentem emisji światła a momentem jego dotarcia do Ziemi. Gdy odległość obiektu jest znana, z tego stosunku można obliczyć, z jaką prędkością rozszerza się wszechświat. Dotyczy to prędkości, która jest dana dzięki stałej Hubble'a, ale też zmiany prędkości, czyli przyspieszenia bądź spowolnienia. W jaki sposób jednak można ustalić odległość między galaktykami? Najbardziej

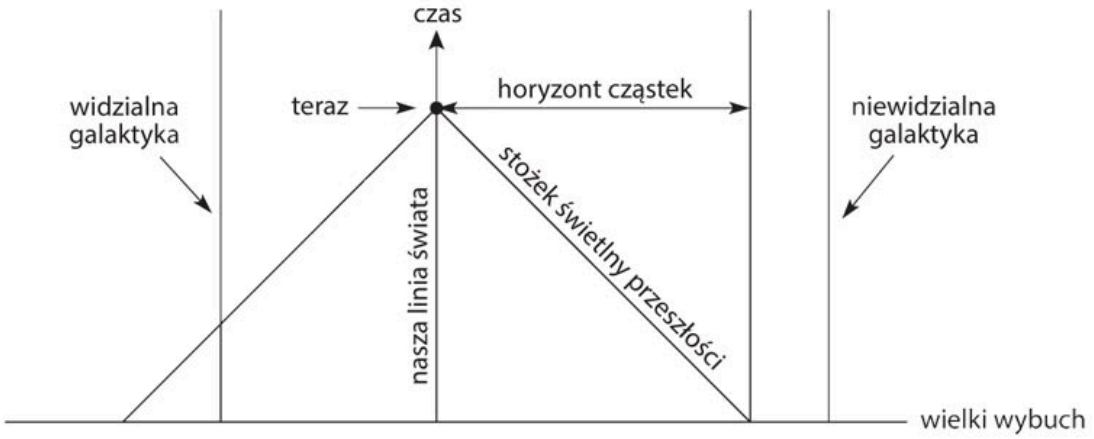
oczywista metoda polega na tym, aby wyszukać w galaktykach poszczególne obiekty, które świecą własnym światłem. Dzięki porównaniu obserwowanej jasności można ustalić ich odległość. W dalszej części niniejszej książki zapoznamy się z takimi obiektami.

Czy wiemy, jak duży jest wszechświat? Niestety nie możemy odpowiedzieć na to pytanie na podstawie obserwacji. Nie wiemy nawet, czy wszechświat jest skończenie, czy też nieskończenie wielki. Łatwiej natomiast odpowiedzieć na pytanie o wielkość widzialnego wszechświata. Jeżeli wszechświat powstał przed skończonym odcinkiem czasu, to światło oddalonych obiektów jest w drodze dopiero od pewnego skończonego czasu. Możemy zatem dostrzec tylko te obiekty, których światło dotarło do nas w tym skończonym czasie. Granica, do której jest to możliwe, nazywana jest *horyzontem*. Odpowiada on horyzontowi na otwartym morzu, który definiuje granicę naszych możliwości obserwacji. W kosmologii używa się dokładniejszego określenia, czyli tak zwanego „horyzontu cząstek". Nazwa ta ma podłoże historyczne i być może powinna brzmieć „horyzont linii świata", ponieważ odnosi się do danej linii świata w czasoprzestrzeni, a w tym przypadku do naszej własnej linii świata. Horyzont cząstek jest więc powierzchnią kuli w trójwymiarowej przestrzeni, która oddziela obiekty obecnie dla nas widzialne od tych obecnie dla nas niewidzialnych.

Aby móc dokładniej określić horyzont cząstek, musimy ponownie ustalić stożek świetlny przeszłości dla teraźniejszego punktu w czasie, czyli drogę promieni świetlnych, które biegną przez czasoprzestrzeń i aktualnie do nas docierają. „Stożek" taki przedstawiono już na rysunku 4.1. Jest on zakrzywiony, gdyż wszechświat rozszerza się, a zatem w przeszłości był mniejszy. Podczas rozważań na temat horyzontów pomocny może okazać się odpowiedni diagram. Chcemy tak dobrać współrzędną przestrzeni, aby linie świata galaktyk umocowanych w jednym punkcie przestrzeni były liniami prostopadłymi. Współrzędna czasu powinna być zaznaczona w taki sposób, aby stożki świetlne przebiegały jak w płaskiej przestrzeni, czyli po

prostych z nachyleniem 45° do osi (mówi się wtedy o „czasie konforemnym”). Rysunek 4.4 ilustruje powyższy opis.

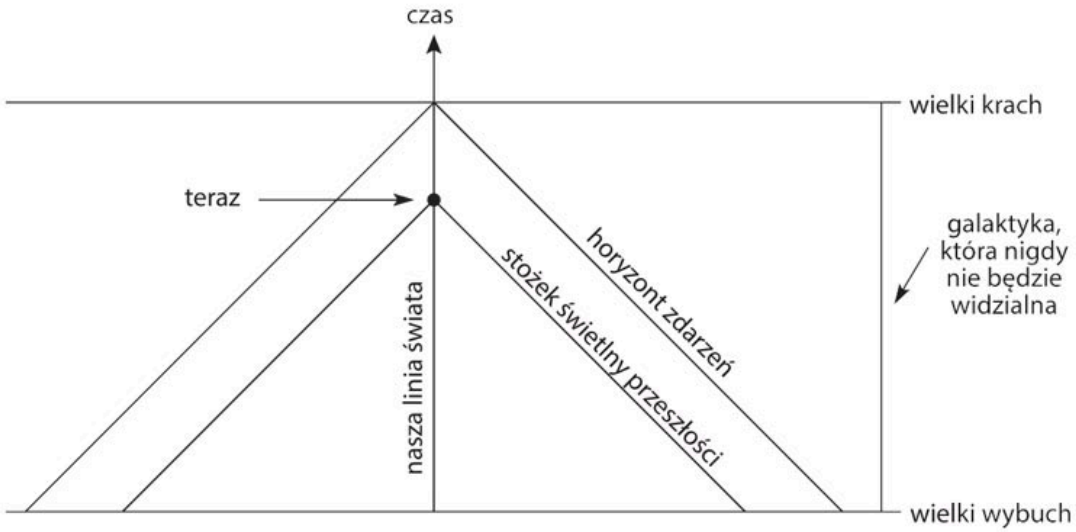
Oś pozioma odpowiada wielkiemu wybuchowi. Z tego względu możemy obserwować tylko te galaktyki, których linie świata przecinają stożek świetlny. Z biegiem czasu stożek świetlny przemieszcza się do góry i linie świata nowych galaktyk wchodzą w nasze pole widzenia. Sam horyzont cząstek pędzi przez galaktyki z prędkością światła. Na rysunku widzimy, że galaktyka, która raz pojawiła się wewnątrz horyzontu cząstek, zawsze już będzie się w nim znajdować i będzie ją można obserwować, dopóki będzie istnieć. Na podstawie obserwacji można wyliczyć wielkość horyzontu cząstek, wynoszącą nieco ponad 14 gigaparseków (Gpc), przy czym 1 Gpc to 1000 Mpc.



4.4. Horyzont cząstek podaje odległość, w ramach której galaktyki są obecnie dla nas widzialne. Na rysunku przedstawiono przypadek wyidealizowany, w którym wszystkie galaktyki istnieją od momentu wielkiego wybuchu.

Szczególnie interesujący jest przypadek zamkniętego wszechświata. Jak głęboko można do niego zajrzeć? Czy dzięki promieniowi świetlnemu, który biegnie do tyłu wokół wszechświata, aby potem ponownie znaleźć się z przodu, można zobaczyć tył swojej głowy? Jak się okazuje, nie jest to możliwe w zamkniętym wszechświecie, który się rozszerza, a potem ponownie kurczy. W momencie jego maksymalnego rozszerzenia można

objąć wzrokiem zaledwie jego połowę, a dopiero podczas wielkiego krachu światło mogłoby objąć go w całości, lecz wtedy i tak jest już za późno.



4.5. Horyzont zdarzeń podaje odległość, w ramach której galaktyki podczas istnienia wszechświata mogą być widzialne. Także tutaj założono przypadek wyidealizowany, w którym wszystkie galaktyki istnieją od momentu wielkiego wybuchu.

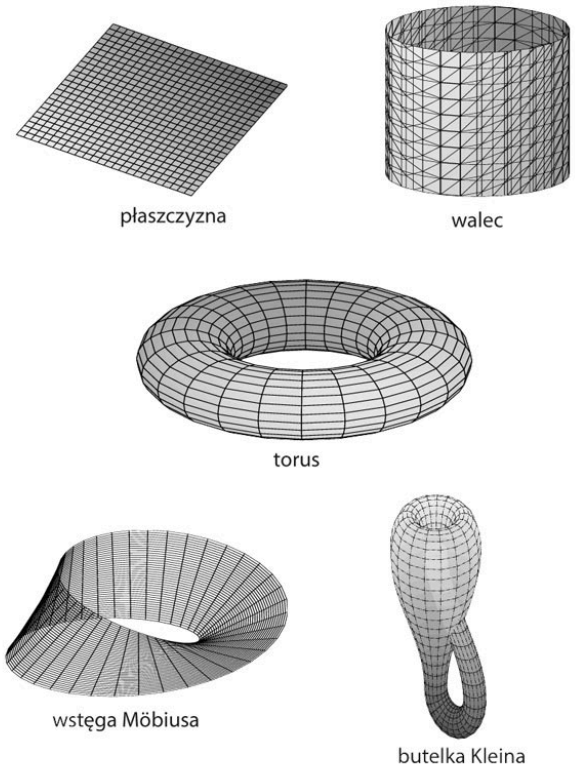
Oprócz horyzontu cząstek istnieje jeszcze jeden ważny horyzont – „horyzont zdarzeń". Poznaliśmy go już w związku z czarnymi dziurami. Tam był on zdefiniowany jako granica, spoza której nic, nawet światło, nie może się wydostać. Analogicznie jest i tutaj. Kosmiczny horyzont zdarzeń oddziela zdarzenia, które kiedyś będą dla nas widzialne, od zdarzeń, które nigdy nie będą dla nas widzialne. To właśnie ilustruje rysunek 4.5. Widać na nim, że istnieje horyzont zdarzeń, gdy wszechświat (w stosunku do konforemnego czasu!) dotrze do swojego końca. W omawianych tu modelach Friedmana-Lemaître'a bez stałej kosmologicznej tylko zamknięty wszechświat, który zmierza do wielkiego krachu, posiada horyzont zdarzeń. Wszechświat nie jest w nich wystarczająco stary, aby pozwolić *wszystkim* obiektom na wysyłanie do nas światła.

Topologia wszechświata

Strona tytułowa czasopisma „Nature” z 9 października 2003 roku przedstawiała dziwny twór składający się z pięciokątów, który przypominał trochę boisko piłkarskie. Postawiono również pytanie: Czy taki jest kształt wszechświata?

Co to znaczy? W poprzednim rozdziale dowiedzieliśmy się, iż zasada kosmologiczna dopuszcza dla przestrzeni trzy rodzaje geometrii, określone przez parametr k. Równania Einsteina definiują jednak tylko lokalną geometrię – nie formułują one żadnych stwierdzeń na temat globalnej struktury przestrzeni. Zagadnieniem tym zajmuje się z kolei topologia (od greckiego słowa „topos”, czyli „miejsce”) – dyscyplina matematyczna, którą nazywano kiedyś *Analysis situs*. Za jej prekursorów uznaje się Leibniza i Eulera. Zajmuje się ona właściwościami przestrzeni, które nie ulegają zmianom mimo ciągłych transformacji. Pod pojęciem ciągłych transformacji rozumie się dowolne deformacje, poza rozrywaniem i zamykaniem pustych przestrzeni. Wspomniana powyżej lokalna geometria dopuszcza istnienie różnych topologii, a więc przestrzeni, które nie mogą ulec rozszczepieniu w wyniku samego ich deformowania („zagniatania”). Przestrzenie o $k = 0$ nazywamy płaskimi, o $k = 1$ – sferycznymi, natomiast o $k = -1$ – hiperbolicznymi.

Z uwagi na zasadę kosmologiczną przedmiotem naszego zainteresowania są wyłącznie przestrzenie o stałej krzywiźnie. W trzech wymiarach sytuacja staje się bardzo skomplikowana, dlatego warto najpierw przyjrzeć się dwóm wymiarom przestrzeni. Ograniczmy się do przestrzeni, które są lokalnie euklidesowe, to znaczy takie, w przypadku których lokalnie obowiązują prawa geometrii euklidesowej. Otrzymujemy zatem pięć możliwości przedstawionych na rysunku 4.6.



4.6. Pięć możliwych przestrzeni płaskich w dwóch wymiarach

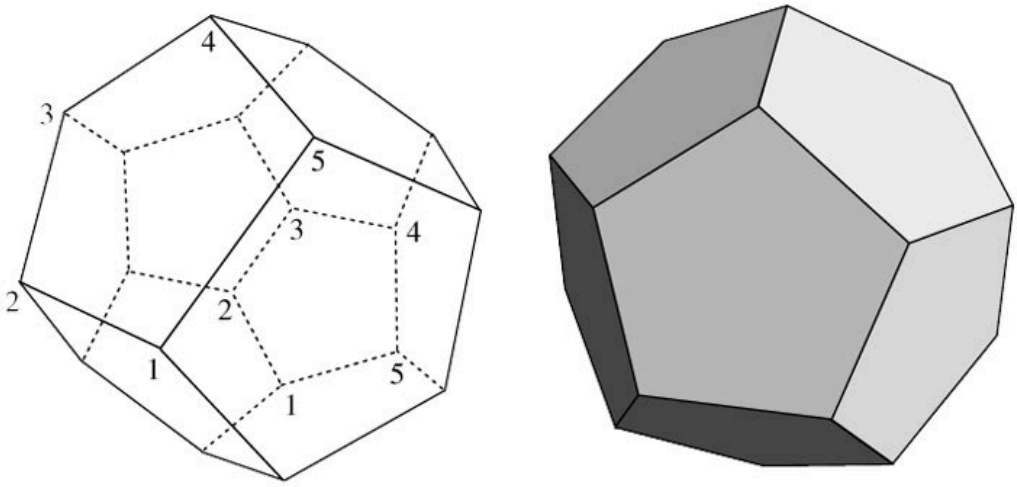
Najprostszym przypadkiem jest zwykła dwuwymiarowa płaszczyzna o nieskończonej długości, kolejnym – walec, który można otrzymać, sklejając ze sobą dwie krawędzie kartki papieru. W ten sposób pierwotnie różne punkty będą identyfikowane wspólnie. Po sklejeniu ze sobą dwóch przeciwległych krawędzi powstaje zamknięta powierzchnia, czyli torus. Na koniec pozostały dwa bardziej skomplikowane przypadki. W wyniku sklejenia dwóch przeciwległych krawędzi – po uprzednim ich skręceniu o 180° – powstaje tak zwana wstęga Möbiusa. Powierzchnia ta posiada tylko jedną stronę: trójwymiarowa mrówka, maszerując w jednym kierunku, dotarłaby do punktu wyjścia od „spodniej strony". Wstęga Möbiusa jest również „nieorientowalna", co oznacza, że dwuwymiarowa mrówka po okrążeniu wstęgi Möbiusa dotarłaby do swojego punktu początkowego w formie lustrzanego odbicia. Niezwykłe, prawda?! Najdziwniejszym przypadkiem jest jednak butelka Kleina, nazwana tak na

cześć matematyka Felixa Kleina (1849–1925). Powstaje ona w wyniku sklejenia ze sobą dwóch przeciwległych krawędzi kartki papieru w sposób standardowy oraz dwóch pozostałych skręconych o 180°. Butelka Kleina łączy niejako właściwości torusa i wstęgi Möbiusa. Jest to nieorientowalna powierzchnia zamknięta. W przestrzeni trójwymiarowej butelkę Kleina można narysować wyłącznie przy wykorzystaniu samoprzecięć (rys. 4.6). Jej powierzchnia jest jednak całkowicie regularna. Z wszystkich przedstawionych wariantów tylko ta powierzchnia jest jednospójna: tylko w jej przypadku możliwe jest zawiązanie kokardy w jednym punkcie. Analogicznie do tych przypadków geometrii euklidesowej istnieje również skończona liczba powierzchni o negatywnej i pozytywnej krzywiźnie (choć tych ostatnich jest bardzo niewiele).

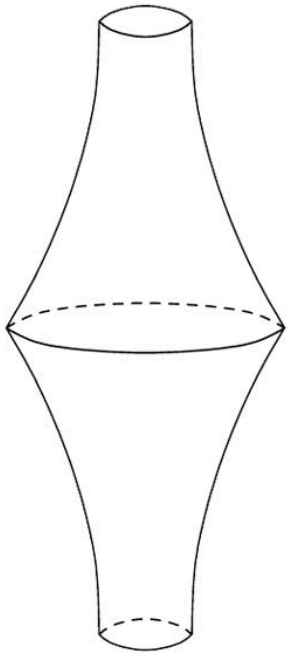
W tej chwili przedmiotem naszego zainteresowania są trójwymiarowe przestrzenie o stałej krzywiźnie. W poprzednim rozdziale poznaliśmy trzy takie przypadki: przestrzeń płaską, sferę i pseudosferę. Są to jedyne przestrzenie jednospójne. Tego, że sfera jest jedyną skończoną trójwymiarową powierzchnią jednospójną, dowiedziono dopiero na początku XXI wieku. Wiąże się to ze słynną hipotezą Poincarégo, za której potwierdzenie rosyjskiemu matematykowi Grigorijowi Perelmanowi przyznano w 2006 roku cenny Medal Fieldsa.

W przypadku rezygnacji z wymogu jednospójności przed topologią otwierają się prawie nieograniczone możliwości. Jedną z nich są geometrie sferyczne, które cieszą się obecnie dużą popularnością. Dowodem na to jest model Einsteina z 1917 roku. Przestrzenie, które nie są jednospójne, można otrzymać, podobnie jak w dwuwymiarowym przypadku, poprzez identyfikację punktów. Na przykład poprzez identyfikację (czyli sklejenie ze sobą) przeciwległych punktów w trójwymiarowej sferze powstaje tak zwana przestrzeń rzutowa lub eliptyczna. Przez niektórych kosmologów, na przykład Arthura Eddingtona, została ona uznana za bardziej naturalną niż sfera trójwymiarowa, ponieważ liczba punktów w obrębie sfery zostaje zmniejszona o połowę, co pozwala uniknąć „nienaturalnego podwojenia” sfery. Podczas gdy przestrzeń rzutowa w kosmologii ma znaczenie przede

wszystkim ze względów historycznych, inna przestrzeń sferyczna cieszy się większym zainteresowaniem: dwunastościenna przestrzeń Poincarégo. To właśnie ta przestrzeń, która zdobiła wymienioną na wstępie stronę tytułową czasopisma „Nature". Dwunastościan foremny poznaliśmy już wcześniej jako jedną z pięciu brył platońskich – składa się on z dwunastu pięciokątów foremnych. Po obróceniu każdego z nich o 36° i zidentyfikowaniu (sklejeniu) z przeciwległym powstaje dwunastościenny wszechświat Poincarégo (rys. 4.7). Jego objętość jest 120 razy mniejsza od sfery. W artykule zamieszczonym w czasopiśmie „Nature" rozważano tę przestrzeń jako możliwy model wszechświata. To taka współczesna analogia do eteru Arystotelesa.



4.7. Dwunastościenna przestrzeń Poincarégo powstaje w wyniku sklejenia przeciwległych pięciokątów



4.8. Pseudosfera

Jakie są inne możliwe krzywizny? Najprostszym przypadkiem w zakresie geometrii hiperbolicznej jest wymieniona już powyżej pseudosfera. Jej dwuwymiarową krewną pokazuje rysunek 4.8.

Ilustracja dowodzi, że pseudosfery jako całości nie można swobodnie osadzić w przestrzeni trójwymiarowej (w środku znajduje się krawędź). Gładka powierzchnia siodła, przedstawiana często w książkach popularnonaukowych jako model przestrzeni hiperbolicznej, nie wygląda na jednorodną, bo taka nie jest – jej krzywizna jest zróżnicowana w różnych punktach. Kwestia osadzenia w przestrzeni płaskiej jest jednak drugorzędna i interesuje nas wyłącznie ze względu na swoją obrazowość – liczy się tylko geometria wewnętrzna przestrzeni. W rzeczywistości każdą przestrzeń (matematycy mówią w tym przypadku o rozmaitości riemannowskiej) można osadzić w przestrzeni płaskiej o wyższych wymiarach – w latach pięćdziesiątych dowiódł tego amerykański matematyk John Nash. Na ogół liczba wymiarów przestrzeni, w której osadzona jest inna przestrzeń, jest znacznie wyższa od wymiaru tej

drugiej. Wspomnianego powyżej dwuwymiarowego torusa (zob. rys. 4.6), charakteryzującego się geometrią płaską, również nie można osadzić płasko w przestrzeni trójwymiarowej – przypominający go donut nie jest płaski, lecz zakrzywiony. Istnieje jednak możliwość płaskiego osadzenia torusa w płaskiej przestrzeni czterowymiarowej.

Z trzech przestrzeni jednospójnych: sfery, pseudosfery i przestrzeni płaskiej tylko sfera ma skończoną objętość. Jak zatem przedstawia się kwestia objętości w przypadku różnych przestrzeni o stałej krzywiźnie? Geometria sferyczna odznacza się tym, że jej objętość jest zawsze skończona, natomiast w przypadku geometrii płaskiej i hiperbolicznej może ona być zarówno skończona, jak i nieskończona. Przykładem zamkniętej przestrzeni trójwymiarowej jest trójwymiarowy torus, krewny wymienionego powyżej torusa o jednym dodatkowym wymiarze. Zamknięte przestrzenie hiperboliczne są mniej znane, choć jest ich nieskończenie wiele. Jak dowiódł amerykański matematyk William Thurston, większość zamkniętych modeli trójwymiarowych dopuszcza geometrię hiperboliczną.

Jak więc w rzeczywistości wygląda nasz wszechświat? Czy na podstawie obserwacji można stwierdzić, czy żyjemy w sferze, wszechświecie dwunastościennym czy w innej przestrzeni? Już Karl Schwarzschild w 1900 roku rozważał możliwość, że nasze życie toczy się w zamkniętym wszechświecie o skomplikowanej topologii – oczywiście jego ówczesne badania dotyczyły wyłącznie naszej Drogi Mlecznej. Z uwagi na to, że przestrzenie te mogą powstawać w wyniku identyfikacji (sklejania) przeciwległych powierzchni, można w pewnym sensie wyjść z jednej strony i automatycznie wejść z drugiej. Promienie świetlne pochodzące od oddalonych obiektów mogą tam dotrzeć do nas wieloma drogami, a my bylibyśmy w stanie zobaczyć wiele obrazów tego samego obiektu. Dodatkowo, w przypadku gdyby przestrzeń była nieorientowalna (jak butelka Kleina), moglibyśmy przemierzać wszechświat, wracając w formie lustrzanego odbicia do naszego punktu wyjścia – zwariowane, prawda? Schwarzschild wywnioskował, że skala, w której zachodzą

możliwe powtórzenia, musi być z całą pewnością większa niż galaktyka, ponieważ jej obrazy wielokrotne nie są dostrzegalne. W cytowanej na wstępie pracy zamieszczonej w czasopiśmie „Nature" przytoczono obserwacje, które rzekomo pasują do dwunastościennego kształtu wszechświata. Obserwacje te dotyczą słabych anizotropii kosmicznego promieniowania tła, o których jeszcze wspomnimy. Inni autorzy nie mogli potwierdzić tych rewelacji, a zatem pytanie o topologię wszechświata nadal pozostaje bez odpowiedzi. W swojej pracy Schwarzschild nawiązuje do wielu możliwych rozwiązań tej kwestii. Według niego każdy z nas żyje poniekąd w geometrycznej bajkowej krainie. Piękne w tym jest to, że nikt nie wie, czy ostatecznie bajka ta nie okaże się rzeczywistością.

Stała kosmologiczna i ciemna energia

Stała kosmologiczna, oznaczana literą Λ (lambda), ma za sobą bogatą historię. Raz istniała, a raz nie, ale ogólnie nigdy nie cieszyła się dużą popularnością. Jak już wspominaliśmy, Einstein wprowadził ją w 1917 roku, aby w ten sposób opisać statyczny model wszechświata, zgodny z zasadą Macha. Po odkryciu przesunięć galaktyk ku czerwieni odstąpił od niej. W 1923 roku na pocztówce do Hermanna Weyla napisał, że skoro nie ma nawet pseudostatycznego świata, to nie ma też stałej kosmologicznej. Według doniesień George'a Gamowa Einstein uznał nawet wprowadzenie Λ za swoją największą pomyłkę. Jednak zarówno Friedman, jak i Lemaître, rozważając równania na jednorodność i izotropowość wszechświata, przewidzieli warianty uwzględniające i pomijające Λ, zatem kategorycznie nie wykluczali jej istnienia.

Jaki wpływ ma przyjęcie termu Λ na rozwój wszechświata? Λ o wartości dodatniej powoduje powstanie grawitacji odpychającej, w przeciwieństwie do przyciągającej natury zwykłej materii. To właśnie ta właściwość w 1917 roku umożliwiła Einsteinowi skonstruowanie statycznego modelu wszechświata, w którym siła odpychająca Λ stanowi przeciwwagę dla przyciągania materii. Powodem odpychania jest fakt, że Λ formalnie odpowiada ujemnemu ciśnieniu, czyli napięciu. W teorii

Einsteina ciśnienie, a nie tylko energia (masa), jak u Newtona, jest źródłem pola grawitacyjnego. Gęstość energii, odpowiadająca termowi Λ, jest dodatnia i stała, niezależnie od wielkości wszechświata ma ona więc zawsze tę samą wartość. W dziedzinie termodynamiki (część 3) mówi się również, że praca, która jest konieczna, aby zwiększyć objętość wszechświata o określoną wartość, jest równa wartości, o którą zwiększy się energia wszechświata przy stałej gęstości energii. W przeciwieństwie do tego pozostaje zwykła materia, która ulega rozcieńczeniu wraz z rozszerzaniem się wszechświata i w związku z tym prowadzi do obniżenia gęstości energii. Ze względu na stałą gęstość energii termu Λ od pewnej wielkości wszechświata stała kosmologiczna będzie zatem zawsze dominować. Tym samym z punktu widzenia losów wszechświata kwestia stałej kosmologicznej ma kluczowe znaczenie. Term Λ o wartości ujemnej odpowiada dodatkowemu przyciąganiu, które w przypadku dużego wszechświata jest zawsze dominujące.

Podczas dyskusji nad modelami Friedmana-Lemaître'a bez Λ widzieliśmy, że istnieje niezaprzeczalny związek między geometrią wszechświata (zarówno płaską, jak i sferyczną czy hiperboliczną) oraz losem wszechświata (zarówno rozszerzającego się na zawsze, jak i ponownie zapadającego). Po uwzględnieniu Λ związku tego nie dostrzegliśmy. Tylko w przypadku Λ o wartości ujemnej sytuacja jest jasna: niezależnie od geometrii wszechświat taki ostatecznie zawsze będzie się zapadał, aż do zniszczenia spowodowanego wielkim krachem. W przypadku geometrii płaskiej ($k = 0$) i hiperbolicznej ($k = -1$) należy pamiętać również o tym, że wszechświat, także w przypadku obecności stałej kosmologicznej o wartości dodatniej, na zawsze się rozszerza – już występująca wieczna ekspansja zostaje zatem jeszcze wzmocniona w wyniku dodatkowego odpychania. Sytuacja komplikuje się w przypadku geometrii sferycznej ($k = 1$). Wówczas wielkość Λ jest decydująca w kontekście tego, czy wszechświat na zawsze się rozszerza, czy ostatecznie zaprzestaje ekspansji i zaczyna się zapadać: gdy Λ jest dość duża, odpychanie przeważa i wszechświat rozszerza się na zawsze. Jeśli Λ jest mniejsza od wartości krytycznej, wtedy zwycięża wpływ materii,

wszechświat zaprzestaje ekspansji i zaczyna się zapadać. Dla tego przypadku istnieje nawet rozwiązanie równań, które opisuje wszechświat pozbawiony osobliwości: początkowo nieskończenie wielki wszechświat zaczyna się kurczyć, po czym osiąga promień minimalny, aby w końcu znów rozszerzać się w nieskończoność.

Istnienie krytycznej wartości Λ daje możliwość dowolnego „spowolnienia" ekspansji wszechświata. Zbliżając w myślach wartość stałej kosmologicznej od wartości lekko podwyższonej do krytycznej, przedłuża się okres, w którym wielkość wszechświata niewiele się zmienia – w pewnym sensie jest to więc stan znajdujący się pomiędzy rozszerzaniem a ponownym kurczeniem się wszechświata, poprzedzający jego dalszą ekspansję. (Osiągnięcie wartości krytycznej jest równoznaczne ze statycznym modelem wszechświata według Einsteina). Kiedyś ten pseudostatyczny model pozwalał rozwiązać pozorne sprzeczności między obserwacjami a modelami bez Λ. Po pierwsze, myślano wtedy, że istnieją gromady kuliste starsze niż świat, którego wiek wyliczono na podstawie modeli bez Λ. W modelu pseudostatycznym można było dowolnie zawyżyć wiek świata, dzięki czemu wszystkie sprzeczności znikały. Po drugie, w jednym z przesunięć ku czerwieni podejrzewano występowanie niespotykanego nagromadzenia kwazarów, które można byłoby zrozumieć, gdyby wszechświat od pewnego czasu znajdował się w fazie tworzenia tych obiektów. Z modelu pseudostatycznego zrezygnowano, gdy obserwacje te udało się wyjaśnić również w modelu tradycyjnym.

Co na podstawie obecnych obserwacji wiemy o wartości Λ? Zgodnie z powyższym warunkiem koniecznym w przypadku obiektów znajdujących się między sobą w odległościach kosmicznych jest związek między absolutną wielkością gwiazdową a przesunięciem ku czerwieni. Do tego celu doskonale nadaje się pewna określona klasa supernowych oznaczanych jako typ Ia. Powstają one, gdy w układzie dwóch gwiazd, składającym się z białego karła i zwykłej gwiazdy, dochodzi do przepływu materii ze zwykłej gwiazdy do białego karła. W ten sposób zwiększa się możliwa masa maksymalna białego karła (ustalona na podstawie teorii

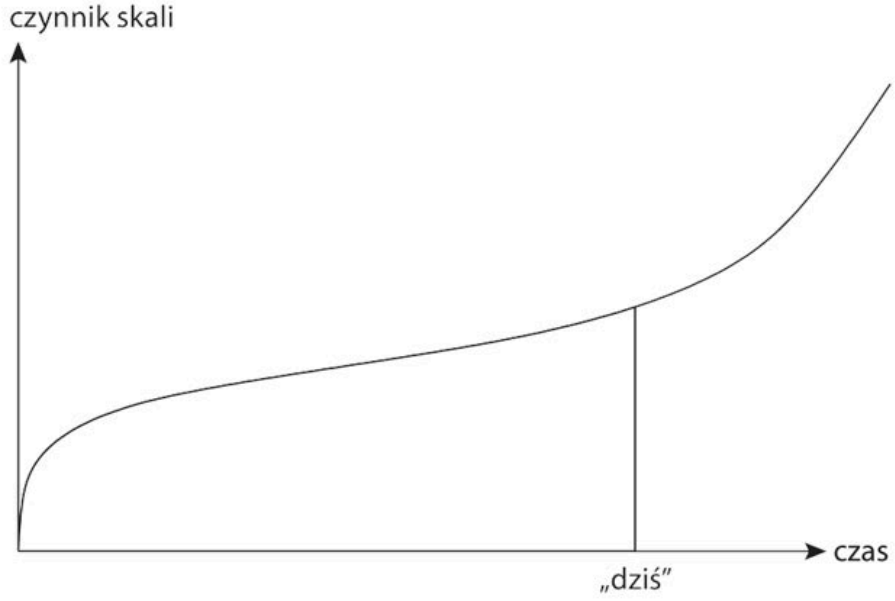
względności), co może doprowadzić do jego eksplozji i powstania supernowej o na tyle dużej jasności, że jej blask jest silniejszy od poświaty galaktyki, do której należy. Ze względu na to, że fizyka tego wybuchu jest zawsze taka sama, można spodziewać się również mniej więcej tej samej jasności – obserwacje widm i przepływu światła w czasie tylko to potwierdzają. Z uwagi na to, że supernowe mają zawsze taką samą jasność, obserwowana wielkość gwiazdowa jest zatem miarą absolutnej wielkości gwiazdowej.

W przypadku wszechświata ze stałą kosmologiczną supernowe wydają się słabsze w porównaniu z wszechświatem bez Λ, ponieważ odległość między nimi jest większa. Obserwacje to potwierdzają! W 1998 roku dwie grupy badaczy ze Stanów Zjednoczonych ogłosiły to po raz pierwszy, a dalsze obserwacje tylko tego dowodzą. Wszechświat znajduje się obecnie w fazie przyspieszonej ekspansji. Rysunek 4.9 przedstawia ewolucję czynnika skali w czasie: wszechświat rozszerza się najpierw umiarkowanie, a po przekroczeniu punktu przegięcia rozszerza się w przyspieszonym tempie. Do przegięcia doszło, gdy stała gęstość energii zawarta w Λ nie przekroczyła jeszcze spadającej gęstości materii. Zaobserwowane supernowe zajmują obszar na wykresie, który zaczyna się jeszcze w fazie umiarkowanego wzrostu i sięga praktycznie do dzisiaj, natomiast bardziej oddalone gwiazdy znajdują się w jeszcze wcześniejszej fazie ekspansji.

Na podstawie modelu, którego dotyczą obserwacje, można również wywnioskować, że nasz wszechświat liczy sobie około 14 miliardów lat. Ponadto można obliczyć odległość od granicy części wszechświata objętej obserwacjami: wynosi ona obecnie wspomniane już powyżej około 14 Gpc. Wprawdzie jest to dystans, którego nie jesteśmy jeszcze w stanie pokonać, jednak chodzi przede wszystkim o przedstawienie aktualnej odległości od najdalszych widocznych nam obiektów – w stanie, w którym były, gdy wszechświat był jeszcze młodszy i mniejszy.

Na rysunku 4.9 można dostrzec jeszcze jedno: do przejścia od umiarkowanej do przyspieszonej ekspansji doszło przed kilkoma miliardami lat, czyli z kosmologicznego punktu widzenia stosunkowo

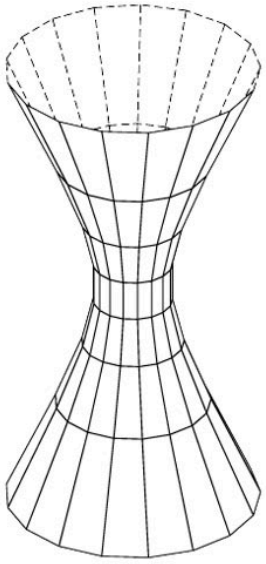
niedawno. Dlaczego nie wcześniej albo później? Innymi słowy: dlaczego zbiegło się to w czasie z powstaniem życia? Przypadek czy wręcz przeciwnie? W tym kontekście często mówi się o problemie koincydencji w kosmologii. Być może można się z nim zmierzyć wyłącznie przy użyciu zasady antropicznej (część 6).



4.9. Ewolucja wielkości naszego wszechświata w czasie (na podstawie czynnika skali); obecnie znajdujemy się już w fazie przyspieszonego wzrostu.

Nasz wszechświat zawiera oczywiście materię, warto jednak przyjrzeć się dokładniej również wszechświatowi, który nie ma materii, a wyłącznie stałą kosmologiczną. Rozwiązanie równań Einsteina nazwano na cześć jego odkrywcy, wymienionego już powyżej astronoma de Sittera, przestrzenią de Sittera (a dokładniej: czasoprzestrzenią). Odgrywa ono szczególnie istotną rolę w rozważaniach nad bardzo młodym wszechświatem (zob. poniżej). Jak można przedstawić tę przestrzeń? Okazuje się, że można ją przedstawić jako czterowymiarową hiperboloidę osadzoną w pięciowymiarowej płaskiej przestrzeni Minkowskiego. Po usunięciu dwóch wymiarów otrzymujemy figurę ukazaną na rysunku 4.10. Hiperboloidę otrzymujemy poprzez obrócenie hiperboli o jedną oś. Na

rysunku przestrzeń w określonym czasie jest po prostu okręgiem. Po ponownym dodaniu dwóch uprzednio pominiętych wymiarów powstaje trójwymiarowa sfera. Czasoprzestrzeń de Sittera można sobie wyobrazić jako sferę, która kurczy się, a po osiągnięciu wartości minimalnej zaczyna się ponownie rozszerzać. Nie ma ona żadnej osobliwości, a zatem w jej przypadku nie można również mówić o wielkim wybuchu.



4.10. Dwuwymiarowy model przestrzeni de Sittera

Dotychczas zawsze zakładaliśmy, że Λ ma wartość dodatnią – sugerują to także wszelkie obserwacje. Oczywiście istnieje też dokładne rozwiązanie równań Einsteina dla stałej kosmologicznej o wartości negatywnej. Nazwano je trochę bez polotu, bo przestrzenią anty-de Sittera. I tę przestrzeń można przedstawić za pomocą hiperboloidy, przy czym jednak w odróżnieniu od rysunku 4.10 kierunki czasu i przestrzeni są zamienione. Przestrzeń anty-de Sittera odgrywa pewną rolę w teorii strun, do której wrócimy jeszcze w kolejnej części tej książki.

Co to wszystko ma wspólnego ze stałą kosmologiczną? Jak podkreślał przede wszystkim rosyjski astrofizyk Jakow Zeldowicz (1914–1987) w 1968 roku, Λ jest bezpośrednio powiązana z energią próżni. W ramach

kwantowej teorii pola energia próżni jest czymś całkowicie naturalnym. Przykładem jej występowania może być efekt Casimira, w wyniku którego dochodzi do powstania siły przyciągania między dwiema nienaładowanymi płytkami umieszczonymi w próżni (część 2). Po pominięciu grawitacji istotne stają się właściwie tylko różnice energii. Dopiero grawitacja nadaje energiom właściwe znaczenie – wynika to ze słynnego wzoru Einsteina $E = mc^2$, według którego każda energia jest równoważna masie, a każda masa wytwarza pole grawitacyjne. Skąd zatem pochodzi energia próżni, wyrażona jako stała kosmologiczna? Jeśli weźmiemy pod uwagę energie próżni występujące w kwantowej teorii pola (nazywane też energiami punktu zerowego), to otrzymamy Λ o wartości przekraczającej tę pochodzącą z obserwacji o wiele rzędów wielkości. Prawdopodobnie jest to jedyny taki przypadek w fizyce, gdy wyniki obserwacji i oczekiwania oparte na założeniach teoretycznych tak dalece od siebie odbiegają. Fakt ten zadziwił również Pauliego. Podobno zastanawiał się on nad oddziaływaniem grawitacyjnym energii punktu zerowego pola elektromagnetycznego. Dla zabawy obliczył wynikającą z tego stałą kosmologiczną i stwierdził, że taki statyczny model wszechświata według Einsteina miałby promień rzędu zaledwie 30 kilometrów. W nawiązaniu do otrzymanych wyników Pauli powiedział sarkastycznie, że taki wszechświat nie objąłby nawet Księżyca.

Być może w przyrodzie istnieje ukryta symetria, która powoduje, że wszystkie energie punktu zerowego sumują się do zera. Pomijając fakt, że do tej pory nie odkryto takiej symetrii, należałoby wtedy spodziewać się Λ, która znika w całości. Jest jednak inaczej: Λ nie znika, ale jej wartość jest wbrew oczekiwaniom bardzo niska. Niektórzy naukowcy uznają ten problem za największą zagadkę we współczesnej fizyce.

Być może suma wszystkich energii punktu zerowego rzeczywiście znika, a zaobserwowana reszta jest pozostałością czegoś innego? Mogłaby na przykład istnieć jakaś nowa forma energii, która tylko imitowałaby stałą kosmologiczną. Taką nową formę energii nazywa się *ciemną energią*. Nie należy jej mylić z ciemną materią. W odróżnieniu od niej oznacza ona

taką formę energii, która rozprzestrzenia się równomiernie i nie skleja się. Ponadto, podobnie jak Λ, musi mieć ujemne ciśnienie. Dobrym kandydatem mogłoby być pole kwantowe o nowej formie wzajemnego oddziaływania, którego (poza jego następstwami jako ciemnej energii) nikt jeszcze nie zaobserwował. Wyrazem tej niewiedzy jest nazwa tego nowego pola określanego najczęściej mianem „kwintesencji", w nawiązaniu do pojęcia eteru wprowadzonego przez Arystotelesa.

Skąd bierze się kwintesencja? W znanych i zbadanych teoriach fizyki cząstek elementarnych pole takie nie występuje. W hipotetycznych rozszerzeniach modelu standardowego obecnych jest za to wiele nowych pól, które są rozważane w kontekście ciemnej energii, jednak żadne z nich zdaje się nie mieć do tego odpowiednich predyspozycji. W związku z powyższym krąży wiele różnych modeli, które są następnie porównywane z obserwacjami supernowych. To, czy dzięki temu uda się doprowadzić do rozwiązania, pozostaje jeszcze kwestią otwartą. W przeciwieństwie do samego termu Λ pole kwintesencji rozwija się w czasie. Dzięki temu istnieje szansa na rozwiązanie wspomnianego powyżej problemu koincydencji.

Od około pięciu miliardów lat dominuje ciemna energia, podczas gdy wcześniej pierwszeństwo należało do ciemnej materii. Pole kwintesencji może mieć taką naturę, która połączy je z kosmologicznym rozwojem materii i promieniowania w taki sposób, że jego dominacja, przynajmniej w przypadku wielu warunków początkowych, przypadnie mniej więcej na ten sam moment. Kwestią otwartą pozostaje jednak nadal, czy stała kosmologiczna bądź ciemna energia zostaną zamknięte w stałych ramach, czy też pozostaną otwarte poprzez swobodny dobór odpowiednich warunków początkowych. W przypadku, gdyby doszło do tego ostatniego, należałoby rzeczywiście bardziej pochylić się nad kwestią koincydencji. Być może jedyną drogą rzeczywiście okaże się zasada antropiczna. Zobaczymy.

Przyszłość wszechświata

Czy można przewidzieć przyszłość wszechświata? Oczywiście byłoby zbyt śmiałym przedsięwzięciem uczynić to na wieczność (czy też choćby tylko na daleką przyszłość). W tym celu zbyt słabo znamy współczesną budowę wszechświata, jego różnorodne formy materii oraz możliwe procesy przemian. Ponadto mamy niedostateczną wiedzę na temat przestrzennych niejednorodności.

Jeśli jednak chodzi o niedaleką przyszłość, to możliwe jest sformułowanie całkiem rozsądnych prognoz. Najprostszą (a zarazem najnudniejszą) jest sytuacja, gdy obecne przyspieszenie spowodowane jest wyłącznie stałą kosmologiczną Λ, a nie zależną w czasie ciemną energią. Można wówczas pokazać, że czasoprzestrzeń przynajmniej lokalnie zbliży się do omawianej powyżej czasoprzestrzeni de Sittera (rys. 4.10). Przypomina to trochę wspomniane w części 1 „łysienie" czarnych dziur, co w tym miejscu przekłada się na to, że wszystkie dodatkowe stopnie swobody („włosy") z biegiem czasu przerzedzą się, aż w końcu zostanie sama energia próżni. Wszechświat rozszerzałby się wtedy zawsze ze stałym przyspieszeniem. Z uwagi na to przyspieszenie inne galaktyki będą oddalać się od nas szybciej niż w przypadku modelu bez Λ. Z tego powodu od niektórych galaktyk odgradzałaby nas granica, której nie bylibyśmy w stanie przekroczyć fizycznie (za pomocą statku kosmicznego) ani radiowo (po wysłaniu odpowiedniego sygnału). Obliczenia dowodzą, że obecnie możemy zobaczyć bardzo wiele galaktyk znajdujących się za tą granicą, a zatem tych, które wprawdzie widzimy, ale z którymi nigdy nie będziemy mogli nawiązać kontaktu. Z biegiem czasu z naszego pola widzenia znikną wszystkie galaktyki, które nie są powiązane grawitacyjnie z naszą Grupą Lokalną. Podobnie jak w przypadku zbliżania się gwiazdy do horyzontu zdarzeń czarnej dziury, obiekty te doświadczą coraz większego przesunięcia ku czerwieni – w czasoprzestrzeni de Sittera rzeczywiście istnieje również horyzont zdarzeń. Wprawdzie nigdy nie będzie nam dane zaobserwować, jak galaktyka opuszcza horyzont, jednak spadek natężenia, idący w parze ze zwiększającym się przesunięciem ku

czerwieni, spowoduje, że galaktyka stanie się niewidoczna. Symulacje komputerowe wykazały, że za 100 miliardów lat w takim dosłownym wszechświecie Λ widzielibyśmy właściwie tylko naszą Drogę Mleczną i Galaktykę Andromedy, które następnie w wyniku działania przyciągania złączyłyby się, tworząc jedną galaktykę. Już po około 30 miliardach lat, a więc okresie mniej więcej dwukrotnie dłuższym, niż liczy sobie obecnie nasz wszechświat, struktury (galaktyki, gromady galaktyk) przestałyby się dalej rozwijać, lecz z uwagi na dominującą ekspansję wszechświata zostałyby w pewnym sensie zamrożone. Do połączenia Drogi Mlecznej i Galaktyki Andromedy w jedną ogromną supergalaktykę doszłoby już wcześniej. Przed przyszłymi obserwatorami, gdyby mieli takowi istnieć, rozpostarłby się niesamowity widok, przedstawiony na rysunku 4.2, skrywający wiele odległych galaktyk i ograniczony do własnego kosmicznego podwórka.

Życia na Ziemi nie będzie już jednak na długo przed tymi odległymi czasami. W przyszłości Słońce znacznie „przytyje", stając się czerwonym olbrzymem, który prawdopodobnie połknie Ziemię. Jak wynika z obliczeń, dojdzie do tego za około 7,6 miliarda lat, ale już za około miliard lat powiększone Słońce doprowadzi do wzrostu temperatury, który uniemożliwi dalsze życie na Ziemi.

Nie wiemy, czy za przyspieszenie ekspansji wszechświata rzeczywiście odpowiedzialny jest term Λ, czy może raczej zmieniająca się w czasie ciemna energia. W takim wypadku istniałaby również ewentualność, że ciemna energia ponownie osłabnie, a wszechświat zatrzyma swoją ekspansję i zacznie się kurczyć aż do zniszczenia spowodowanego wielkim krachem. Jednak z ostrożnych szacunków opierających się na aktualnych obserwacjach wynika, że faza ekspansji potrwa jeszcze przynajmniej przez 20 miliardów lat.

W zależności od dokładnej natury ciemnej energii możliwe są różne scenariusze co do przyszłości naszego wszechświata. Jeden z nich określa się mianem „wielkiego rozdarcia". Scenariusz ten zakłada, że wszechświat w skończonym czasie nieskończenie się powiększy, co ostatecznie

doprowadzi do swoistego „rozdarcia”. Inną ewentualnością jest „wielkie zatrzymanie”. W ramach tego scenariusza w przyszłości wszechświat wyhamuje nieskończenie szybko, a potem nie zmieni już swojej wielkości. Następnie, gdy gęstość materii zniknie, ciśnienie materii nieskończenie wzrośnie, a zatem dojdzie do wystąpienia osobliwości. Na podstawie ostrożnych szacunków można powiedzieć, że minie jeszcze co najmniej 20 miliardów lat, zanim powstaną takie osobliwości jak wielkie rozdarcie, o ile w ogóle to nastąpi. Wielkie zatrzymanie może natomiast stać się faktem w dużo bliższej przyszłości. Na podstawie obecnych obserwacji obliczono, że może do tego dojść „już” za około 10 milionów (nie miliardów) lat. Jeśli nawet liczba ta nie spędza nikomu snu z powiek, to scenariusz ten byłby brzemienny w skutkach – przecież znaczyłoby to, że wszechświat większość swojego życia ma już za sobą. To tyle, jeśli chodzi o spekulacje co do naszej kosmicznej przyszłości. A jak wyglądała przeszłość naszego wszechświata?

Przeszłość wszechświata

W jaki sposób możemy dowiedzieć się czegoś na temat przeszłości naszego wszechświata? W tym celu musimy stworzyć harmonijny obraz między teorią a obserwacją. Obowiązującą teorią jest ogólna teoria względności Einsteina uzupełniona współczesnymi teoriami fizyki cząstek elementarnych i fizyki jądrowej. Dzięki tym środkom możemy prześledzić historię wszechświata do czasów, gdy jego istnienie zamykało się w zaledwie jednej sekundzie. By cofnąć się jeszcze bardziej, musimy nawiązać do teorii kwantowej grawitacji, którą poznamy dokładniej dopiero w kolejnej części tej książki.

Dzięki obserwacjom zapoznaliśmy się już z kilkoma ważnymi przykładami. I tak na podstawie wielkości gwiazdowych supernowych w odległościach kosmicznych mogliśmy wnioskować, że aktualnie jesteśmy świadkami przyspieszonej ekspansji wszechświata. Jednoznaczny obraz uzyskamy jednak dopiero po uwzględnieniu innych obserwacji. Na podstawie rozmieszczenia galaktyk i gromad galaktyk oraz ich dynamiki

można wyciągać wnioski przede wszystkim co do zawartości materii. Nawet najbardziej konserwatywni astronomowie są zgodni co do tego, że parametr Ω dla materii wynosi znacznie mniej niż jeden, a co za tym idzie – gęstość materii jest (o około 30 procent) mniejsza od gęstości krytycznej wszechświata.

Z obserwacji kosmicznego promieniowania tła wynika jeden ważny czynnik. Jak już wspomnieliśmy, promieniowanie to jest w wysokim stopniu izotropowe i osiąga temperaturę prawie trzech kelwinów. Jednak w rzeczywistości, gdy dokładniej się temu przyjrzeć, można zaobserwować niewielkie wahania temperatury. Po pierwsze, można mówić o znanej już od dłuższego czasu „anizotropii dipolowej": w jednym kierunku temperatura jest lekko podwyższona, a w przeciwnym lekko obniżona – podwyższenie bądź obniżenie temperatury wyrażone jest w tym wypadku w milikelwinach (jedna tysięczna kelwina). Jest to przykład występowania efektu Dopplera, który odzwierciedla ruch względem kosmicznego układu odniesienia, ustalonego za pośrednictwem promieniowania tła. Na ruch ten składa się obrót Ziemi wokół Słońca, Słońca wokół centrum Galaktyki oraz ruch Drogi Mlecznej i Grupy Lokalnej. Pomiary mówią o prędkości rzędu 600 kilometrów na sekundę względem kosmicznego układu odniesienia.

4.11. Fluktuacje w kosmicznym promieniowaniu tła przedstawione na podstawie obserwacji satelity WMAP po pięciu latach analizowania danych

Po pominięciu anizotropii dipolowej wahania temperatury są jeszcze mniejsze (wyraża się je w mikrokelwinach, czyli jednej milionowej kelwina), a ich podział ma znaczenie wyłącznie statystyczne. Pierwszego pomiaru tych wahań dokonał satelita COBE w 1992 roku. Nowsze wyniki dostarczane są z wielu różnych źródeł – ostatnie pochodzą z sondy WMAP, wyniesionej w przestrzeń kosmiczną przez amerykańską agencję kosmiczną NASA w 2001 roku (skrót pochodzi od „Wilkinson Microwave Anisotropy Probe"). Na 2008 rok europejska agencja kosmiczna ESA planuje z kolei start satelity PLANCK, której zadaniem będzie dokładniejsze zbadanie wspomnianej powyżej anizotropii promieniowania. Rysunek 4.11 przedstawia obraz nieba uwzględniający anizotropie promieniowania tła wychwycone przez satelitę WMAP.

Czemu służy to ogromne przedsięwzięcie? Dane na temat wahań temperatury zawierają wiele informacji o naszym wszechświecie. Na ich podstawie można zatem ustalić szereg parametrów kosmologicznych. Zgodnie z powyższym można na przykład stwierdzić, że parametr Ω musi mieć wartość mniej więcej równą jeden. Wszechświat znajduje się więc dość blisko gęstości krytycznej. Skoro materia zajmuje tylko około 30 procent, to brakujące 70 procent musi należeć do ciemnej energii. Z połączenia tych pomiarów z danymi o supernowych powstaje harmonijny obraz gęstości materii i energii we wszechświecie.

Większa część z tych 30 procent zajmowanych przez materię należy do ciemnej materii, której natura nie została jeszcze dobrze zbadana. Materia widzialna, czyli głównie zwykłe gwiazdy, odpowiada za zaledwie pięć promili. Czym jest zatem pozostała część? Istnieje pewność co do tego, że składa się ona przynajmniej z dwóch różnych elementów. Jeden z nich musi mieć naturę barionową, a drugi – naturę niebarionową. Barionami nazywa się cząstki składające się z trzech kwarków. Najważniejszymi ich przykładami są cząstki, z których my sami jesteśmy zbudowani, czyli protony i neutrony, będące składnikami jądra atomowego. Zakłada się, że część barionowa ma wartość Ω na poziomie pięciu procent. Przynależna ciemna materia mogłaby składać się z zimnych gwiazd lub ciemnych

chmur gazowych. Co z pozostałymi 25 procentami? Ta część ciemnej materii nie może składać się z materii, którą znamy, dlatego określa się ją mianem „egzotycznej ciemnej materii". Być może są to nienaładowane cząstki pochodzące z obszerniejszej teorii cząstek elementarnych, których bezpośrednio jeszcze nie zaobserwowano. W tym kontekście często mówi się o tak zwanym neutralino, czyli najlżejszej z cząstek, których istnienie przewiduje teoria supersymetrii (porównaj z częściami 2 oraz 5). Do jej powstania mogło dojść we wczesnej fazie rozwoju wszechświata. Od tej pory jej liczebność we wszechświecie mogła zwiększać się w sposób właściwie niezakłócony. Akceleratory cząstek, takie jak LHC w ośrodku CERN, zajmują się ich poszukiwaniem. Jak dotąd są to jednak zwykłe, nawet jeśli nie całkowicie bezpodstawne, spekulacje.

Co możemy powiedzieć o młodości wszechświata? Kluczowych informacji w tym względzie dostarcza nam znów promieniowanie tła. W obecnym wszechświecie jest ono słabe i pod kątem produkcji energii pozostaje daleko w tyle za materią. W związku z tym mówi się o fazie wszechświata zdominowanej przez materię. Nie zawsze jednak tak było. Wraz z ekspansją wszechświata energia promieniowania maleje szybciej niż energia materii – oprócz zwykłego rozcieńczenia energii dochodzi również do ubytków spowodowanych ochłodzeniem. Jeśli spojrzeć wstecz na historię wszechświata, można zauważyć, że energia promieniowania była kiedyś wyższa od energii materii – wszechświat był wtedy zatem zdominowany przez promieniowanie. Prawdopodobnie był to okres, gdy wszechświat nie miał jeszcze nawet 50 000 lat.

Sytuacja we wszechświecie zdominowanym przez promieniowanie jest w pewnym sensie prosta. Stan ten można scharakteryzować na podstawie jego temperatury, która poprzez równania Friedmana i Lemaître'a łączy się z wiekiem świata. Termodynamika przewiduje również zależność między temperaturą a energią. Ponadto trzeba jeszcze wiedzieć, z jakimi cząstkami i oddziaływaniami mamy do czynienia. Zgodnie ze słynnym $E = mc^2$ Einsteina masa m cząstki łączy się z energią E. W przypadku, gdy temperatura wszechświata jest tak wysoka, że odpowiadająca jej energia

jest wyższa od tego *E*, może dojść do powstania cząstek promieniowania, które ją zrównoważą. Wynika to z kwantowej teorii pola, którą omówiliśmy w części 2 tej książki. Historia wszechświata przedstawia więc różne obszary energii z zakresu fizyki cząstek elementarnych. Z drugiej strony na podstawie naszych doświadczeń wyniesionych z tego działu fizyki jesteśmy w stanie odtworzyć historię wszechświata w warunkach laboratoryjnych. Wszechświat w swoim wczesnym stadium rozwoju działa zatem jak akcelerator cząstek, wytwarzając energię, która wykracza poza nasze ziemskie wyobrażenie.

Spróbujmy teraz odtworzyć historię wszechświata zdominowanego przez promieniowanie, zaczynając od końca. U schyłku fazy zdominowanej przez materię wszechświat składał się głównie z najlżejszych atomów, wolnych elektronów i promieniowania. Przede wszystkim były to jądra atomowe wodoru (czyli protony) oraz helu. Występowały również śladowe ilości litu, berylu i boru. Wszystkie inne metale ciężkie powstały dużo później – w wyniku wybuchu supernowych. Gdy świat liczył sobie niecałą minutę, składniki jądra atomowego, czyli protony i neutrony, były jeszcze niepowiązane. Dopiero potem wszechświat schłodził się na tyle, że protony i neutrony mogły się połączyć, tworząc najlżejsze jądra atomowe. Proces ten określa się mianem *nukleosyntezy*. Obserwacje potwierdzają rozpowszechnienie pierwiastków wynikające z obliczeń, na przykład mniej więcej jedna czwarta zaobserwowanego helu powstała w okresie wczesnego rozwoju wszechświata. Należy także podkreślić, że sukces nukleosyntezy jest kolejnym dowodem na istnienie ogólnej teorii względności. Na podstawie nukleosyntezy oblicza się wspomnianą wyżej zawartość barionową wszechświata, która stanowi około pięciu procent gęstości energii.

Co wydarzyło się przed nukleosyntezą, w pierwszej minucie istnienia świata? W pierwszych sekundach energia wszechświata była wystarczająca, aby z promieniowania wyodrębniły się elektrony i pozytony, z którymi oddziaływały ponadto neutrina. W krótkim czasie doszło do dwóch ważnych wydarzeń. Najpierw neutrina odłączyły się od

innych cząstek i od tej pory rozwijały się niezależnie od reszty wszechświata, a następnie uległy znacznemu ochłodzeniu i podobnie jak kosmiczne promieniowanie tła istnieją również dziś, choć nie ma na to bezpośredniego dowodu.

Po odłączeniu się neutrin energia wszechświata spadła poniżej poziomu masy elektronu, a większość elektronów i pozytonów przeistoczyła się w promieniowanie i tylko niewielka, choć istotna, część elektronów przetrwała, tworząc później wraz z protonami i neutronami atomy stanowiące podstawę naszej egzystencji. W wyniku tej anihilacji wszechświat nieco się zagrzał – temperatura przekroczyła temperaturę odłączonych uprzednio neutrin. Z tego powodu istniejące także dziś neutrinowe promieniowanie tła charakteryzuje się niższą temperaturą niż kosmiczne promieniowanie tła (różnica wynosi około dwóch kelwinów).

Kosmiczne neutrina, podobnie jak fotony kosmicznego promieniowana tła, przenikają współczesny wszechświat i otaczają nas w dużej liczbie. Z uwagi na ich słabe oddziaływanie dostarczenie bezpośredniego dowodu na ich istnienie wydaje się w najbliższym czasie zadaniem niemożliwym do zrealizowania. Pośrednio mamy jednak dowód na istnienie kosmicznego morza neutrin, co w marcu 2008 roku ogłosił zespół należący do misji WMAP. Neutrina uczestniczą bowiem w rozkładzie anizotropii w kosmicznym promieniowaniu tła.

Zarówno protony, jak i neutrony składają się z trzech kwarków. Kwarki są do tego stopnia ze sobą połączone (poprzez cząstki wymienne zwane gluonami), że zazwyczaj nie występują jako cząstki swobodne. Gdy świat liczył sobie jednak jedną tysięczną sekundy, energia wszechświata była tak duża, że kwarki i gluony występowały jako cząstki swobodne. We wczesnym okresie rozwoju wszechświata istniały oczywiście również antycząstki kwarków, czyli antykwarki. Gdy temperatura wszechświata spadła, kwarki i antykwarki przeistoczyły się w promieniowanie. Dlaczego w takim razie kwarki nadal istnieją? Powód jest taki sam jak w przypadku elektronów. Po prostu kwarków było troszkę więcej niż antykwarków – około trzech kwarków więcej na każdy miliard. Ocalała tylko ta niewielka

część antykwarków. Przyroda zachowuje zatem fundamentalną asymetrię między barionami a antybarionami – na nasze szczęście. Przyczyn istnienia tej asymetrii należy prawdopodobnie upatrywać w fundamentalnych teoriach fizyki cząstek elementarnych, które są istotne w przypadku działania bardzo wysokich energii. Szczegóły nadal pozostają jednak dla nas zagadką. Asymetrię barionową można też opisać zgodnie z modelem standardowym fizyki cząstek elementarnych, choć nie jest ona wystarczająca, aby wyjaśnić powstanie asymetrii barionowej we wczesnej fazie istnienia wszechświata.

Teraz wrócimy do czasów nieco późniejszych, gdy wszechświat wkroczył w fazę dominacji materii, czyli około 50 000 lat po wielkim wybuchu. W tym czasie materia była jeszcze zjonizowana, a jądra atomowe oderwane od elektronów. W obu tych przypadkach dochodziło do silnego oddziaływania z promieniowaniem, które z tego powodu jest powiązane z cząstkami. Pod pojęciem materii rozumie się tutaj „zwykłą materię". W rzeczywistości istnieje jeszcze również wspomniana już wcześniej ciemna materia, która odgrywa istotną rolę w procesie tworzenia struktur wszechświata. Podczas gdy zwykła materia jest powiązana z promieniowaniem, ciemna materia z uwagi na własną grawitację może się sklejać – tworzyć zarodki struktur (niejednorodności).

Około 380 000 lat po wielkim wybuchu, gdy w wyniku ekspansji temperatura spadła do około 3000 kelwinów, w wyniku połączenia elektronów z jądrami atomowymi mogło wreszcie dojść do powstania atomów neutralnych. Istniejące już wtedy niejednorodności ciemnej materii zaczęły przyciągać te atomy, co znacznie przyspieszyło proces powstawania struktur wszechświata. Następnie doszło do uwolnienia promieniowania elektromagnetycznego, które od tej pory krąży bez większych przeszkód we wszechświecie. W związku z tym mówi się o tym okresie jako o epoce „odłączenia". To właśnie to uwolnione promieniowanie, które teraz – po schłodzeniu do zaledwie 2,7 kelwina – możemy obserwować pod postacią kosmicznego promieniowania tła. Promieniowanie to jest odzwierciedleniem wszechświata takim, jaki był,

mając 380 000 lat (rys. 4.11). Ważne informacje na temat młodego wszechświata są zawarte w niewielkich wahaniach temperatury tego promieniowania, będących przedmiotem następnego rozdziału tej książki. Epoka odłączenia promieniowania i materii zapowiada nadejście „ciemnej fazy" w historii wszechświata, w której promieniowanie schładza się na długości fali o niewidzialnym zakresie, podczas gdy materia nie zaczyna jeszcze świecić. Dopiero powstanie pierwszych gwiazd i galaktyk – gdy świat liczył sobie około 500 milionów lat – ponownie rozświetliło wszechświat. Te pierwsze galaktyki możemy zobaczyć na zdjęciu zrobionym przez kosmiczny teleskop Hubble'a (rys. 4.2). Do powstania naszego Układu Słonecznego doszło, gdy wszechświat miał około dziewięciu miliardów lat. W okresie tym rozpoczęła się również faza dominacji energii próżni we wszechświecie. Około miliarda lat później proces powstawania galaktyk dobiegł końca – narodził się współczesny nam wszechświat.

Wszechświat inflacyjny

Jak sami mogliśmy się przekonać, istnieją mocne przesłanki ku temu, by uznać, że nasz wszechświat znajduje się obecnie w fazie przyspieszonego wzrostu. Z drugiej strony wiele przemawia też za tym, że wszechświat w bardzo wczesnej fazie swojego rozwoju przechodził już bardzo podobny okres. Wtedy przyczyną powstania przyspieszenia mógł być również jakiś rodzaj energii próżni, działającej jak stała kosmologiczna, która w błyskawicznym tempie z impetem powiększyła wszechświat. Przyjmuje się, że do ekspansji tej doszło między 10^{-36} a 10^{-34} sekundy po wielkim wybuchu, czyli w bardzo wczesnej fazie istnienia wszechświata. Z uwagi na to, że ekspansja wszechświata odbywa się w równie szybkim tempie jak dewaluacja pieniądza w czasach galopującej inflacji (za przykład mogą służyć Niemcy i rok 1923), przyjęły się następujące określenia: „faza inflacji", „wszechświat inflacyjny" czy po prostu „inflacja". Ze względu na wzrastającą dominację stałej kosmologicznej (bądź ciemnej energii) we współczesnym wszechświecie stoimy w pewnym sensie u progu

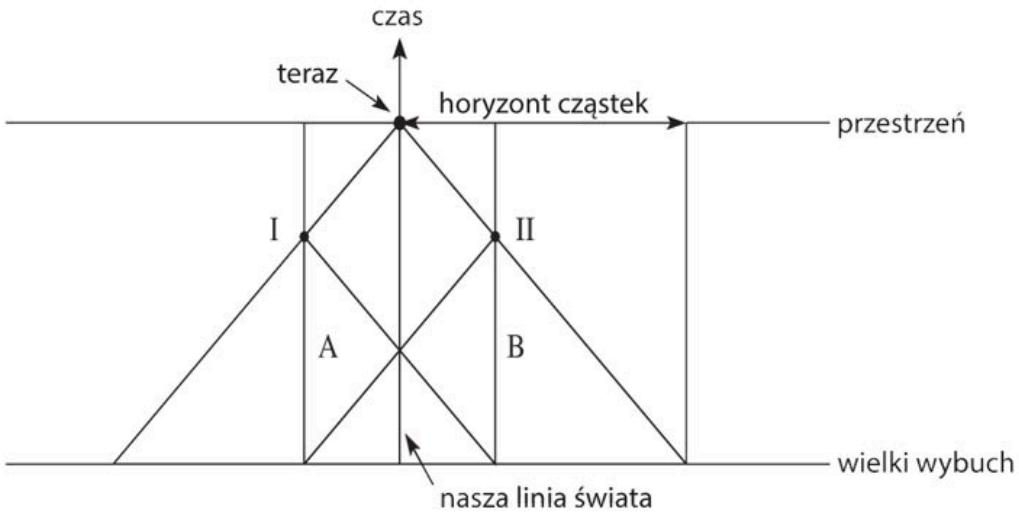
kolejnej fazy inflacji. Pytanie, czy obie te fazy mają to samo podłoże fizyczne, pozostaje nadal otwarte.

Czy są powody ku temu, aby uznać, że nadszedł okres inflacji? Pierwszy model, który odnosił się do tej kwestii, powstał w 1980 roku. Jego autorem był rosyjski fizyk Aleksiej Starobinski. Chciał on opracować model opisujący wszechświat pozbawiony osobliwości początkowej, ponieważ w przypadku istnienia stałej kosmologicznej lub energii próżni warunki niezbędne do poparcia twierdzeń o osobliwościach Hawkinga i Penrose'a nie musiałyby być spełnione. Popularność wszechświata inflacyjnego wzrosła w 1981 roku po opublikowaniu przez amerykańskiego fizyka Alana Gutha pracy, która diametralnie zmieniła obraz współczesnej kosmologii. Guth nie tylko wprowadził termin „inflacja" na określenie przyspieszonej ekspansji wszechświata, ale również zwrócił uwagę na to, że pojęcie to przyczyni się do rozwiązania przynajmniej trzech problemów związanych ze starym modelem wielkiego wybuchu. Jakie to były problemy? Pierwszy z nich jest problemem, który samemu sobie stworzono. Skłonił on jednak Gutha do zajęcia się kwestią inflacji. Teorie spekulatywne z zakresu fizyki cząstek elementarnych, dotyczące ujednolicenia wszystkich oddziaływań niegrawitacyjnych (zob. część 2), przewidują istnienie nowych cząstek oraz monopoli magnetycznych o dużej masie. Jak dotąd nie udało się ich jeszcze zaobserwować, lecz gdyby teorie unifikacji okazały się prawdziwe, to we wczesnym wszechświecie, gdy dominowały potężne energie, doszłoby do powstania wielu monopoli – byłoby ich tak wiele, że również obecnie nie mielibyśmy problemu z ich dostrzeżeniem (co, jak wiemy, nie jest prawdą). Gdy jednak przyjmiemy, że po fazie powstawania monopoli doszło do przyspieszonej ekspansji wszechświata, która spowodowała rozpierzchnięcie się i rozcieńczenie monopoli – w efekcie czego doświadczamy obecnie trudności z ich odnalezieniem – inflacja ukazuje się jako rozwiązanie powstałego na papierze problemu. Z drugiej strony rozważania spekulatywne są potrzebne, aby móc uzasadnić istnienie inflacji, ponieważ zakładają one działanie energii próżni, która jest niezbędna do przyspieszenia wszechświata.

Dużo poważniejsze są dwa pozostałe problemy: problem płaskości i horyzontu. Najpierw kilka słów o problemie płaskości. Przedstawione powyżej obserwacje pokazują, że parametr gęstości Ω oscyluje wokół cyfry jeden, a więc w dużym przybliżeniu powierzchnia wszechświata jest płaska. Dlaczego stanowi to problem? W przypadku, gdy wszechświat nie rozszerza się w tempie przyspieszonym, lecz spowolnionym, parametr Ω z biegiem czasu coraz bardziej oddala się od cyfry jeden – uczą tego równania Friedmana i Lemaître'a. Jeśli zatem wszechświat liczy sobie prawie 14 miliardów lat, a parametr ten nadal jest bliski jedynce, znaczy to, że kiedyś musiał być jej jeszcze bliższy bądź nawet równy jeden. Nie jest to oczywiście żadna sprzeczność logiczna, gdyż wszechświat może być po prostu tak „zaprogramowany", jednak wśród fizyków jest niewielu zwolenników takiego zaprogramowania, gdy nie stoi za nim żaden konkretny powód. Jak zaraz się przekonamy, przyjęcie fazy inflacji rozwiązuje ten problem.

Problem horyzontu ma bardzo podobny charakter. Nie wyraża on również żadnej logicznej sprzeczności starego modelu wielkiego wybuchu, lecz przyprawia jedynie o ból głowy, gdy przyjmie się (jak to się robi w fizyce), że warunki z przeszłości mogą „wyjaśniać" teraźniejszość, kiedy wyjdzie się od swoistej przyczynowości intuicyjnej. Problem ten ma związek z wysoką izotropią promieniowania tła. Promieniowanie to jest pozostałością po gorącej wczesnej fazie rozwoju wszechświata. 380 000 lat po wielkim wybuchu zostało ono odłączone od materii i uwolnione. W momencie, gdy do tego doszło, wiele obszarów promieniowania było jednak od siebie niezależnych – od chwili wielkiego wybuchu nie upłynęło jeszcze dostatecznie wiele czasu, aby wszystkie części tego promieniowania mogły zostać powiązane. Za przykład może służyć promieniowanie działające w przeciwnych kierunkach, które możemy obecnie obserwować na niebie. W chwili odłączenia dzieliło je od siebie 86 milionów lat świetlnych, czyli dużo więcej niż osiągalne dla nas 380 000 lat świetlnych. Skąd zatem ta zadziwiająca izotropia? Należałoby przypuszczać, że doszło do niej w wyniku obecności silnego

oddziaływania. W jaki jednak sposób bez istnienia oddziaływania mogło dojść do osiągnięcia niemal identycznej temperatury?



4.12. Problem horyzontu; diagram odpowiada rysunkowi 4.4

Problem związku przyczynowego wiąże się z horyzontem, czyli granicą obszaru, do którego mogło dotrzeć światło od czasu wielkiego wybuchu. Każdemu przypadkowi istnienia horyzontu, a dokładniej: horyzontu cząstek, towarzyszy również problem horyzontu. Rozważmy sytuację z rysunku 4.12 – podobną do tej z rysunku 4.4. Czas (a dokładniej: równokątna współrzędna czasu) wskazuje do góry, a przestrzeń (a dokładniej: jeden z trzech wymiarów przestrzeni) – w prawo. Linie świata dla obiektów („galaktyk") są liniami pionowymi. Światło porusza się po prostych, które są nachylone względem osi pod kątem 45°. Rozważmy dwie galaktyki A i B, które w stosunku do nas położone są w takiej samej odległości, ale w przeciwnych kierunkach. Obecnie widzimy światło emitowane przez galaktykę A jako zdarzenie I oraz światło emitowane przez galaktykę B jako zdarzenie II. W chwili, gdy światło zostało wyemitowane przez obie galaktyki, nie znajdowały się one jeszcze w żadnym związku przyczynowym (bądź właśnie wówczas znalazły się w takim związku), ponieważ linia świata jednej z nich nie znalazła się w obrębie stożka przeszłości drugiej. W przypadku, gdy

galaktyki wyglądają podobnie, można zadać sobie pytanie, jak to możliwe, skoro nie było wspólnego procesu stworzenia, który byłby wywołany bezpośrednim oddziaływaniem obu tych galaktyk na siebie. W przypadku promieniowania tła mamy do czynienia z promieniowaniem pochodzącym z dwóch przeciwnych kierunków, które zostało wyemitowane znacznie przed zdarzeniem I czy II. Nie istniała zatem możliwość „umówienia się" na jakąś wspólną temperaturę. W związku z tym należy przyjąć, że izotropia promieniowania istniała od początku, a co za tym idzie – nie było żadnego związku przyczynowego. Na tym właśnie polega problem horyzontu.

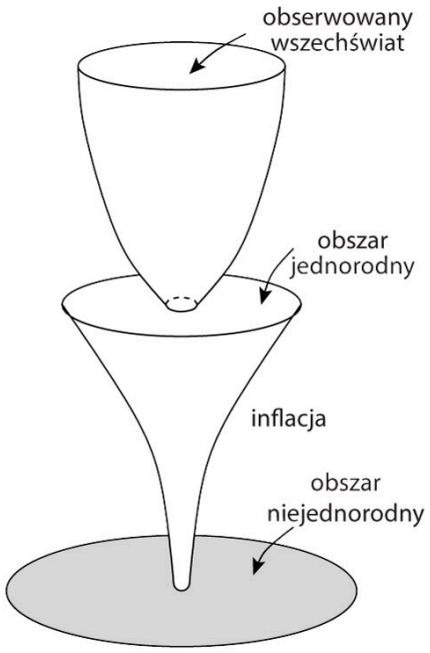
Problem horyzontu ma również inne oblicze. Zgodnie z powszechnym wyobrażeniem galaktyki powstały w wyniku wahań gęstości we wczesnym wszechświecie. Ze względu na problem horyzontu w starym modelu wielkiego wybuchu wahania te nie miały jednak ze sobą żadnej styczności. W tym kontekście wyjaśnienie przyczyny powstania galaktyki nie jest zatem możliwe, co jest oczywiście sytuacją niezadowalającą.

Problemy te można rozwiązać, gdy przyjmie się, że wszechświat w młodości przeszedł fazę inflacji. Za przykład niech posłuży nam fragment przestrzeni, który krótko po wielkim wybuchu jest tak mały, że wszystkie jego punkty znajdują się w związku przyczynowym. Energia w tym okresie jest tak duża, że potrzebne są teorie spekulatywne z zakresu fizyki cząstek elementarnych, aby ją opisać. Takich obszarów było oczywiście więcej. Przyjmijmy dalej, że przynajmniej w kilku tych obszarach dominuje energia próżni, co prowadzi do efektów podobnych do tych, które można zaobserwować w przypadku stałej kosmologicznej – dany obszar bardzo szybko („wykładniczo") się rozrasta. Niewielka część tego powiększonego obszaru przeistacza się później w znany nam obecnie wszechświat. Proces ten można zaobserwować na rysunku 4.13.

Oznaczałoby to, że obecnie widzimy tylko niewielką część wszechświata, a dokładniej tę, która jest w dużej mierze jednorodna. Teraz staje się również jasne, dlaczego inflacja rozwiązuje problem płaskości i horyzontu. Dla dwuwymiarowej istoty stojącej na

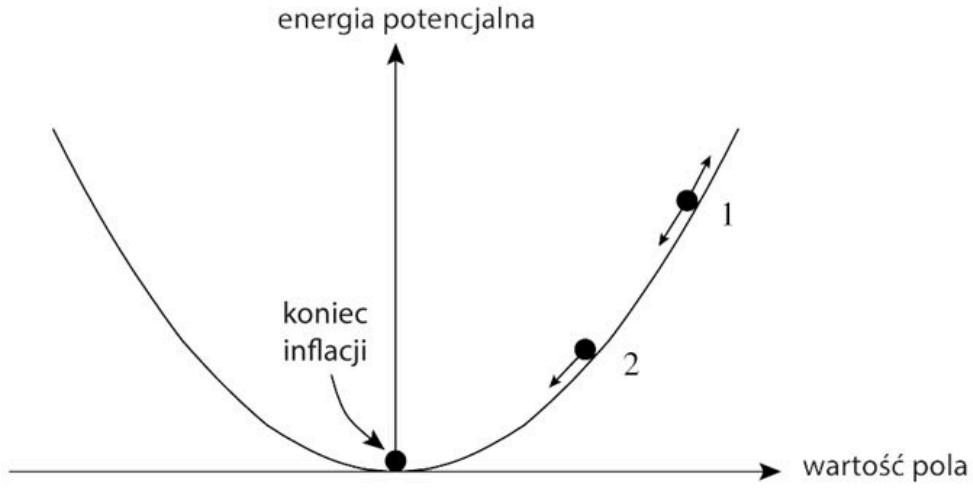
nadmuchanym balonie zawsze będzie on płaski. To samo odnosi się do obserwatora, który stojąc w miejscu, postrzega Ziemię jako płaską powierzchnię. Niezależnie od tego, jak duża jest krzywizna wszechświata w jego odległych punktach, część, którą widzimy, wydaje się nam płaska. Z tego powodu parametr gęstości oscyluje wokół jedynki.

Inflacja rozwiązuje problem horyzontu, ponieważ wszechświat, który możemy zobaczyć, wywodzi się z *jednego* obszaru odznaczającego się spójnością. Regiony, z których dociera do nas kosmiczne promieniowanie tła, były kiedyś ze sobą powiązane i dzięki temu mamy do czynienia z taką samą temperaturą promieniowania. Za pomocą inflacji i związków przyczynowych w zadowalający sposób można wyjaśnić również powstanie i rozwój galaktyk.



4.13. Niewielkie obszary jednorodne w fazie poprzedzającej inflację. Podczas inflacji dochodzi do ich znacznego wzrostu. Z niewielkiej części tego nadmuchanego obszaru powstaje współczesny nam wszechświat, który jest w dużej mierze jednorodny.

To tyle, jeśli chodzi o ogólną ideę inflacji. Czy można ją zrealizować, opierając się na jakimś modelu fizycznym? W tym przypadku sprawy nie wyglądają już tak różowo. Jest wiele modeli, które dają takie możliwości. Na ogół są one skonstruowane w taki sposób, aby przez odpowiedni czas zapewnić warunki potrzebne do powstania inflacji, jednak tylko nieliczne z tych modeli spełniają wymogi praw fizyki. Podobnie jak w przypadku ciemnej energii trzeba posłużyć się jednym bądź kilkoma polami kwantowymi (zwanymi polami inflatonowymi), które posiadają odpowiednią energię potencjalną, aby zasymulować działanie stałej kosmologicznej. Większość z wymienianych obecnie modeli pochodzi od tak zwanej „inflacji chaotycznej", opracowanej w 1983 roku przez rosyjskiego fizyka Andrieja Lindego. Nazwa tej teorii ma nawiązywać do faktu, że powinna ona sprawdzić się w dowolnych, czyli w pewnym sensie „chaotycznych", warunkach początkowych. Początkowo dopuszczalne jest zatem bardzo niejednorodne rozmieszczenie inflatonu i krzywizny przestrzeni, a równowaga termiczna nie jest wymagana. Należy jednak pamiętać, że parametry energii potencjalnej muszą być starannie dobrane, aby w ogóle doszło do powstania inflacji.



4.14. Model inflacji chaotycznej Lindego

Na rysunku 4.14 przedstawiono typowy przykład inflacji chaotycznej. Wykres ten ukazuje stosunek energii potencjalnej pola inflatonowego do wartości pola. Przyjmijmy, że początkowa wartość pola jest wysoka (na rysunku oznaczono ją cyfrą 1). Kwantowa teoria pola uczy nas, że występują duże fluktuacje kwantowe, a potencjał inflatonu pod wpływem niewielkich wahań przypadkowych będzie odpowiednio rósł i malał. Inflaton symuluje energię próżni, o ile wartość pola pozostaje mniej więcej stała. Wszechświat rozszerza się wtedy w przyspieszonym tempie. W przypadku nieco niższych wartości początkowych (pozycja 2) fluktuacje kwantowe są również mniejsze. Pole, jak w lepkiej cieczy, porusza się wolno w dół. W trakcie tego ruchu symuluje ono nadal energię próżni, która jednak stopniowo słabnie. Po opadnięciu na dno pole oscyluje wokół minimalnej wartości w punkcie równowagi, energia próżni zanika, a faza inflacji jest zakończona. Zgodnie z kwantową teorią pola wahania te prowadzą do powstania wielu cząstek, które w wyniku działania energii kinetycznej są w stanie podwyższyć temperaturę wszechświata. Konsekwencją tego jest rozpoczęcie zdominowanej przez promieniowanie fazy wszechświata. Podczas przejścia do tej fazy dochodzi do powstania właściwie każdej materii we wszechświecie. Od tego momentu (czyli około 10^{-34} sekundy po wielkim wybuchu) rozpoczyna się okres ewolucji zgodnie ze starym modelem wielkiego wybuchu i zgodnie z tym, co opisano powyżej. Szczegóły dotyczące wzrostu temperatury nie są jeszcze znane.

W przypadku mniej lub bardziej dowolnych warunków początkowych w ramach inflacji chaotycznej wszystkie wartości pola muszą być przybliżone. Wystąpienie inflacji możliwe jest jednak wyłącznie w obszarach o dostatecznie dużym polu. Obszary te zostają zatem znacznie rozdmuchane i dominują w obrębie swojej całkowitej objętości. Każdy z tych obszarów staje się również jednorodny. Widoczna część naszego wszechświata wywodzi się właśnie z takiego obszaru. Obraz ten z pewnością brzmi dość dziwacznie: wszechświat, którego prawdziwej wielkości nie znamy, składający się z wielu niezwiązanych ze sobą obszarów, z których nigdy nie poznamy więcej niż cząstki jednego z nich.

W tym kontekście mówi się czasami o „multiświecie" składającym się z wielu niezależnych od siebie wszechświatów. Nowe wiadomości z dziedziny grawitacji kwantowej (zob. następna część) w pewnym sensie potwierdzają ten obraz. Ów ogromny wszechświat (czy też multiświat) nie ma ani wspólnego wielkiego wybuchu, ani wspólnego wielkiego krachu, bo przecież nie jest już jednorodny. Niektóre obszary na zawsze się rozrosną, a niektóre – całkiem skurczą.

W obszarach o dużych wartościach początkowych pola inflatonowego (obszar 1 na rys. 4.14) fluktuacje kwantowe w pewnych częściach mogą być tak duże, że potencjał pola jeszcze bardziej wzrośnie. W częściach tych ekspansja będzie kontynuowana. W ich zakresie znajdą się również kolejne części o jeszcze większych wartościach pola itd. Zawsze wystąpią zatem obszary, które będą się rozrastać pod wpływem inflacji. W tym kontekście Linde mówi o niekończącej się samoreprodukcji wszechświata inflacyjnego.

Obraz, jaki dostarcza nam ten scenariusz, jest o wiele bardziej osobliwy od tego oferowanego przez stary model wielkiego wybuchu. Mimowolnie przypomina nam się kastylijski król Alfons Mądry, który powiedział podobno, że mógłby lepiej doradzać Bogu podczas tworzenia wszechświata, gdyby był przy tym obecny – mówiąc to, myślał o mało wyrafinowanym ptolemeuszowskim modelu wszechświata.

Czy taki obraz świata faktycznie odpowiada rzeczywistości? Tego niestety nie wiemy. Dzięki prowadzeniu obserwacji można jednak potwierdzić lub obalić hipotezę istnienia fazy inflacji. Założenie istnienia inflacji pozwala z dużą dokładnością stwierdzić, że przestrzeń jest płaska. Inflacja wskazuje także na pewne właściwości świadczące o niewielkich anizotropiach w promieniowaniu tła. Obserwacje zdają się tylko to potwierdzać. Związek między anizotropiami a inflacją oraz teorią kwantową zostanie wyjaśniony w kolejnym rozdziale. Obecnie próżno szukać poważnej alternatywy dla inflacji. Jak mawiał Linde: inflacja to najlepsza klacz w stajni. Niestety obecnie również jedyna.

Fluktuacje kwantowe a powstawanie struktur wszechświata

Kosmiczne promieniowanie tła przedstawia nam obraz wszechświata 380 000 lat po wielkim wybuchu. Promieniowanie to jest w wysokim stopniu izotropowe. Dociera ono do nas z każdej strony z temperaturą około 2,7 kelwina. Istnieją jednak drobne odstępstwa (anizotropie), wynoszące mniej niż jedną tysięczną kelwina (porównaj z rys. 4.11). Te niewielkie anizotropie stanowią źródło wszelkich niejednorodności we wszechświecie, począwszy od gromad galaktyk i galaktyk, a skończywszy na gwiazdach i planetach. W przypadku idealnie izotropowego promieniowania powstawanie struktur wszechświata nie byłoby możliwe. Wszechświat pozostałby wtedy jednorodny, a przez to niezwykle nudny (nie byłoby jednak wtedy nikogo, kto mógłby się nudzić).

Przyczyną tych niewielkich wahań temperatury są fluktuacje w rozkładzie gęstości materii podczas odłączania fotonów. Nierozłącznie towarzyszą temu fluktuacje w polu grawitacyjnym, a tym samym również w geometrii przestrzennej. Przestrzeń składa się z wielu niewielkich, rozmieszczonych równomiernie dolin i pagórków. Z uwagi na to w przypadku fotonów, które docierają do nas z tych obszarów, dochodzi do omówionego w części 1 przesunięcia ku czerwieni bądź fioletowi. Foton, który dociera do nas z doliny, ma za sobą dłuższą drogę w polu grawitacyjnym i dlatego w porównaniu z fotonem z „równiny” (uśrednionym polem grawitacyjnym) charakteryzuje się przesunięciem ku czerwieni, podczas gdy foton z pagórka odznacza się przesunięciem ku fioletowi. Przesunięcie ku czerwieni odpowiada spadkowi temperatury, a przesunięcie ku fioletowi – wzrostowi temperatury, co prowadzi do powstania czegoś przypominającego patchworkowy dywan przedstawiający niewielkie wahania temperatury (rys. 4.11). Jest to więc, podobnie jak w przypadku systemu nawigacji satelitarnej GPS, efekt ogólnej teorii względności.

Po dokładniejszym zbadaniu różnic temperatury przy mniejszym kącie można zauważyć kilka regularności. W obrębie anizotropii temperaturowej

widoczne są wyraźne wierzchołki (*peaks*), które powstały w wyniku drgań, podobnie jak te pochodzące od fal dźwiękowych, dlatego nazywa się je wierzchołkami akustycznymi. Ich korzenie sięgają czasów, gdy fotony i bariony były jeszcze ze sobą połączone. Podczas gdy ciśnienie fotonów skierowane jest na zewnątrz, aby usunąć anizotropie, bariony, z uwagi na przyciąganie grawitacyjne, kurczą się. Efektem tego są wspólne drgania. W momencie, w którym dochodzi do odłączenia, różne długości fali „cieczy fotonowo-barionowej" zostają zamrożone w ich aktualnym stanie oscylacji. Długości fali, które w tym momencie znajdują się w stanie maksymalnego zagęszczenia lub rozrzedzenia, prowadzą do osiągnięcia maksymalnej wartości anizotropii temperaturowej, ponieważ odłączone fotony zostają poddane maksymalnemu przesunięciu ku czerwieni lub fioletowi. Stąd biorą się wierzchołki akustyczne.

Z widocznego na rysunku 4.11 rozkładu wahań temperatury można odczytać sporo informacji. Jego dokładna forma zależy bowiem od wielu istotnych parametrów kosmologicznych. Precyzyjny pomiar rozkładu temperatury w połączeniu z przytoczonymi wyżej obserwacjami supernowych zapewnia spójny obraz wszechświata, znajdującego się obecnie w fazie przyspieszonej ekspansji, składającego się w dużej mierze z ciemnej materii i ciemnej energii, liczącego około 13,7 miliarda lat i którego przestrzeń jest niemal płaska.

Skąd jednak biorą się niewielkie fluktuacje, w wyniku których doszło do powstania galaktyk? W starym modelu wielkiego wybuchu, który nie uwzględniał inflacji, kwestia ta pozostawała zagadką. Zgodnie z tym, co zostało już powiedziane wcześniej w związku z problemem horyzontu, w starym modelu w momencie odłączenia promieniowania większa część wszechświata od chwili wielkiego wybuchu nie miała możliwości wzajemnego oddziaływania. Równomierność temperatury, jak i niewielkie jej wahania pozostają zatem niewyjaśnione.

Również w odniesieniu do wyjaśnienia fluktuacji przyjęcie istnienia fazy inflacji pełni funkcję koła ratunkowego. Obszar wszechświata, z którego pochodzi promieniowanie tła, znajdował się w związku

przyczynowym w bardzo wczesnej fazie (na początku inflacji). Zakłada się, że w fazie tej nie było jeszcze tego, co obecnie nazywamy materią i promieniowaniem. Poza polem grawitacyjnym, oraz być może innymi nieznanymi polami, istniało jeszcze pole inflatonowe, które było odpowiedzialne za wystąpienie fazy inflacji. Zgodnie z dominującą opinią pole grawitacyjne i inflatonowe występowały w stanie podstawowym, czyli, zgodnie z teorią kwantową, w stanie próżni. W takim wypadku istniałyby tylko nieuniknione fluktuacje kwantowe, przewidziane przez zasadę nieoznaczoności Heisenberga – wirtualne cząstki ciągle niszczone i powstające na nowo. Dlaczego stan próżni? Czy nie mógłby istnieć bardziej skomplikowany stan kwantowy? Na pewno tak, ale są argumenty potwierdzające tę hipotezę. Po pierwsze, jest ona przytaczana w rozważaniach dotyczących kosmologii kwantowej (zob. część 6). Po drugie, stan próżni to stan najprostszy i najbardziej symetryczny, a aspekt estetyczny jest w fizyce nie do przecenienia.

W momencie uwolnienia promieniowania fluktuacje gęstości nie są jednak fluktuacjami kwantowymi. Są to wahania, które można opisać w ramach fizyki klasycznej. Jak dochodzi do powstania tych klasycznych fluktuacji z fluktuacji kwantowych bardzo młodego wszechświata? W ten sposób wracamy do omawianych w części 2 pytań dotyczących podstaw teorii kwantowej. Wtedy dowiedzieliśmy się, jak powstają właściwości klasyczne, a mianowicie poprzez dekoherencję, czyli nieodwracalne i nieuniknione oddziaływanie z otoczeniem. Dekoherencja odpowiada również za to, że we wczesnym wszechświecie z fluktuacji kwantowych (zwanych też pierwotnymi fluktuacjami kwantowymi) powstały klasyczne fluktuacje gęstości. Jak mogło do tego dojść? Skąd wzięło się otoczenie?

Otoczenie jest tylko metaforą stopni swobody, które oddziałują z zainteresowanym układem, same jednak pozostają niedostępne (nieistotne). W wyniku oddziaływania dochodzi do skrzyżowania układu z otoczeniem, co prowadzi do tego, że właściwości z dziedziny mechaniki kwantowej układu zanikają w powiązaniach z otoczeniem – w samym układzie pozostają zatem niewidoczne. Za otoczenie fluktuacji

pierwotnych mogłyby służyć inne pola o nieznanych jeszcze właściwościach bądź też części układu (grawitacja i inflaton), których nie można dostrzec. Na przykład mogłyby to być niewielkie długości fali, które oddziałują z widocznymi długościami fali, ale nie mogą występować samodzielnie. Inną ewentualnością mogłyby być części długości fali, które występują po obu stronach horyzontu i z tego powodu pozostają niewidoczne. W trakcie fazy inflacji ze stanu próżni powstaje skomplikowany, wzbudzony stan kwantowy. Obliczenia pokazują, że w okresie inflacji działa również dekoherencja zamieniająca pierwotne fluktuacje kwantowe we fluktuacje gęstości, które zachowują się w sposób klasyczny i prowadzą do powstawania wahań temperatury w kosmicznym promieniowaniu tła.

Przyjęty pierwotnie wysoce symetryczny stan próżni jest jednorodny. Wpływ koherencji także tego nie zmienia. Stan całego układu jest jednorodnym efektem nałożenia się na siebie (superpozycji) niejednorodnych stanów pojedynczych. Aby zrozumieć, dlaczego możemy zaobserwować klasyczne niejednorodności, musimy znaleźć związek z hipotezami teorii kwantowej. Jeśli dochodzi do kolapsu funkcji falowej, to zamiast superpozycji obserwujemy tylko niejednorodny stan pojedynczy. Jeżeli natomiast pozostaniemy przy hipotezie wielu światów (zwanej też hipotezą Everetta), to otrzymamy obserwatora w wielu wariantach, z których każdy dostrzega wyłącznie niejednorodny stan pojedynczy w obrębie jednorodnej superpozycji. Obserwowany wynik jest za każdym razem taki sam i nie zależy od interpretacji. Zatem również w fizyce wczesnego wszechświata mamy do czynienia z problemem pomiaru teorii kwantowej.

Zgodnie z tym scenariuszem (a lepszego obecnie nie ma) struktury wszechświata, a więc gromady galaktyk i galaktyki, powstały we wczesnym wszechświecie w wyniku działania nieuniknionych fluktuacji kwantowych, przewidzianych przez zasadę nieoznaczoności. Odpowiedni rozkład prawdopodobieństwa w mechanice kwantowej odzwierciedla się

w rozmieszczeniu tych struktur w przestrzeni. Tylko kosmos kwantowy pozwoli nam zrozumieć, skąd pochodzimy.

5. Grawitacja kwantowa

W ostatecznym rozrachunku wiedza
opiera się na założeniach.

Ludwig Wittgenstein

Dlaczego grawitacja nie jest klasyczna?

W lipcu 1955 roku w Bernie spotkało się 89 naukowców, aby świętować pięćdziesiąte „urodziny” szczególnej teorii względności. Przy tej okazji chciano również podsumować postępy w pracy nad szczególną i ogólną teorią względności oraz przedstawić najbardziej palące problemy. Sam Einstein nie mógł uczestniczyć w tym spotkaniu – zmarł trzy miesiące wcześniej w Princeton.

Konferencji przewodniczył Wolfgang Pauli, przez wielu uznawany za swoistego następcę Einsteina. W swojej mowie końcowej Pauli nawiązał do granic naszej wiedzy i kwestii kwantyzacji pola, opisując trudności z tym związane. Chodzi naturalnie o kwantyzację pola grawitacyjnego i opracowanie teorii, która w sposób spójny połączyłaby teorię kwantową z teorią względności, czyli o teorię *grawitacji kwantowej*. Nie należy przez to rozumieć teorii kwantowej dla cząstek i pól *w* zewnętrznej (w znacznym stopniu zakrzywionej) czasoprzestrzeni, lecz teorię kwantową dla samej czasoprzestrzeni, ponieważ geometria czasoprzestrzeni jest przecież właśnie manifestacją grawitacji. W grawitacji kwantowej przestrzeń i czas nie powinny już dłużej odgrywać roli nieruchomego tła, lecz mieć w równym stopniu wpływ na działanie cząstek i pól. Właśnie ta dynamiczna rola przestrzeni i czasu była dla Pauliego główną trudnością w kontekście kwantyzacji grawitacji.

Peter Bergmann (1915–2002) – fizyk niemieckiego pochodzenia, który wyemigrował do Stanów Zjednoczonych, uznawany za eksperta w dziedzinie kwantyzacji grawitacji – w swoim przemówieniu w Bernie powiedział, że na pytanie, czy potrafi dokonać kwantyzacji grawitacji, musi niestety nadal odpowiadać: „Jeszcze nie". Zapytany jednak, czy w ogóle jest to możliwe, stwierdził, że według niego można szczerze odpowiedzieć: „Najprawdopodobniej tak". Jednak nawet teraz, ponad 50 lat po tej konferencji, sformułowanie teorii kwantowej grawitacji uznaje się nadal za nierozwiązany problem. Czym jest to spowodowane? I dlaczego problem ten jest tak ważny, że wiele pokoleń fizyków teoretycznych troszczyło się i nadal troszczy o jego rozwiązanie?

Ważnym, jeśli nie najważniejszym, powodem jest wiara w unifikację wszystkich oddziaływań. Odpowiada to stanowisku zajmowanemu przez redukcjonistów, zgodnie z którym fizyka ogranicza się do zaledwie kilku teorii i oddziaływań, ostatecznie sprowadzając się do jednej „teorii dla wszystkiego" bądź *teorii wszystkiego*, będącej, przynajmniej w zasadzie, źródłem występowania poszczególnych oddziaływań i zjawisk. Historia nauki zdaje się potwierdzać to stanowisko. Pierwotnie niezależne od siebie teorie z działu optyki, magnetyzmu i elektryczności okazały się później różnymi zagadnieniami jednej teorii elektromagnetyzmu. Nowym przykładem skutecznego redukcjonizmu jest połączenie sił słabych i elektromagnetycznych w oddziaływanie elektrosłabe, które obok oddziaływania silnego jest częścią modelu standardowego fizyki cząstek elementarnych.

Z jakiego powodu stworzenie zunifikowanego obrazu grawitacji miałoby się nie udać? Przecież grawitacja łączy się z każdym rodzajem energii, ponieważ każda energia odpowiada masie, a masa wytwarza pole grawitacyjne. Z uwagi na to, że teorie kwantowe z powodzeniem opisują już wszystkie inne oddziaływania, z grawitacją nie powinno albo raczej nie może być przecież inaczej.

Uwzględnienie grawitacji mogłoby okazać się korzystne również dla innych teorii kwantowych. Zgodnie z tym, co przedstawiono w części 2,

teorie te zawierają wiele nieskończoności, które trzeba mozolnie opakowywać w parametry dające się dowieść jedynie za pomocą doświadczeń, a nie obliczeń. Nieskończoności te (*dywergencje*) występują dlatego, że operuje się na dowolnie krótkich długościach i zgodnie z zasadą nieoznaczoności tym samym na dowolnie wysokich pędach i energiach. Kwantyzacja przestrzeni i czasu mogłaby zatem okazać się pomocna. Głównymi orędownikami tej teorii byli Heisenberg i Pauli. W kolejnych rozdziałach niniejszej książki przekonamy się, w jakim stopniu uda się zrealizować te idee.

Zapewne będą również tacy, których nie przekonają przytoczone powyżej powody. Czy trzeba szukać nowych teorii, skoro wszystkie znane nam wyniki obserwacji i eksperymentów da się wyjaśnić za pomocą już sprawdzonych teorii? Otóż trzeba. Te sprawdzone już teorie nie mówią bowiem wszystkiego. Poznaliśmy już wcześniej twierdzenia o osobliwościach ogólnej teorii względności: w ogólnych warunkach można pokazać, że są granice w czasoprzestrzeni, poza które teoria Einsteina nie sięga. Granice te istnieją jednak w sferach istotnych z punktu widzenia fizyki! Wnętrze czarnych dziur oraz ich faza parowania należą do tego grona tak samo jak początek wszechświata. Do zrozumienia także i tych kwestii niezbędna jest bardziej ogólna teoria. Z uwagi na rolę, jaką w przeszłości odegrała mechanika kwantowa, która pozwoliła rozwiązać analogiczne problemy klasycznej elektrodynamiki (jak na przykład niestabilność atomów), uznaje się, że to grawitacja kwantowa będzie mogła obyć się bez osobliwości i stworzyć jednoznaczny obraz dla początku wszechświata i rozwoju czarnych dziur.

Nie można jeszcze uznać, że fundamenty fizyki są już wzniesione. O fakcie tym świadczy chyba najbardziej „problem czasu". Ogólna teoria względności i teoria kwantowa przedstawiają diametralnie różne wyobrażenia czasu. Mechanika kwantowa przejęła całkowicie czas absolutny Newtona. Czas jest w tym wypadku parametrem zewnętrznym, niezależnym od wydarzeń fizycznych, który jest jednak niezbędny do interpretacji teorii kwantowej. W ten sposób funkcje falowe rozwijają się

w czasie, przy czym całe prawdopodobieństwo pozostaje zachowane, czyli jest niezależne od czasu. Również szczególna teoria względności niewiele zmienia w tej kwestii. Wprawdzie przestrzeń i czas łączą się, tworząc czterowymiarową czasoprzestrzeń, ale pozostawia ona tylko miejsce dla dynamiki, która nie ma wpływu na strukturę czasoprzestrzeni. Dopiero ogólna teoria względności uwalnia czasoprzestrzeń z tych więzów i podporządkowuje ją dynamicznym równaniom fizycznym – równaniom Einsteina. W ten sposób czas jako część czasoprzestrzeni stał się dynamiczny. Skoro czas nie może być jednocześnie absolutny (jak w teorii kwantowej) i dynamiczny (jak w teorii względności), oznacza to, że współczesna fizyka jest niekompletna. Oczekuje się, że połączona teoria grawitacji kwantowej zharmonizuje pojęcie czasu. Jak mogłoby do tego dojść, przekonamy się później.

W części 2 mogliśmy zobaczyć, że wszystko wskazuje na uniwersalność teorii kwantowej. W szczególności zasada superpozycji zdaje się potwierdzać, o ile doświadczenia na to pozwalają. Oczywiście uniwersalności tej nie da się dowieść matematycznie, a jedynie potwierdzić empirycznie. Jeśli jednak założy się jej istnienie, to oznacza to, że pole grawitacyjne także *musi* mieć naturę kwantową. Superpozycja jednego obiektu w różnych miejscach jest w rozumieniu mechaniki kwantowej również zawsze superpozycją jego masy, a tym samym wytworzonego przez niego pola grawitacyjnego. Zatem i pole grawitacyjne spełnia wymogi zasady superpozycji w rozumieniu mechaniki kwantowej, chyba że założymy, iż w takim wypadku teoria kwantowa nie obowiązuje. Byłby to wtedy przykład wystąpienia dynamicznego kolapsu funkcji falowej.

Jeśli pole grawitacyjne jest rzeczywiście polem kwantowym, to ma to poważne skutki dla kosmosu jako całości – staje się on wtedy kosmosem kwantowym. Z uwagi na to, że grawitacja jest jedynym oddziaływaniem, które dominuje w dużych skalach, fundamenty wszechświata można poznać wyłącznie za pomocą grawitacji kwantowej. Sposób

funkcjonowania kosmologii kwantowej oraz wnioski, które ze sobą niesie, zostaną opisane w następnej części tej książki.

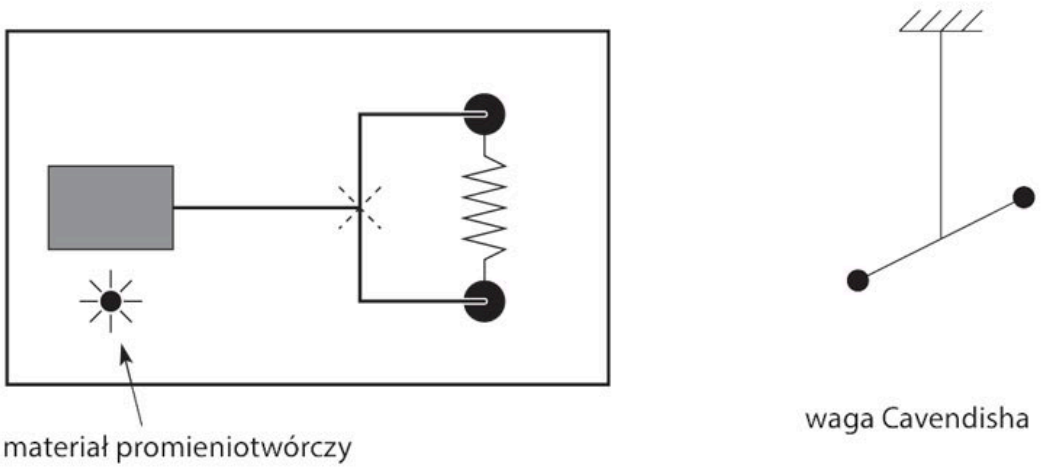
Jednym z pionierów poszukiwań grawitacji kwantowej był Matwiej Bronstein (1906–1938). Na początku lat trzydziestych XX wieku należał on, wraz z Lwem Landauem, George'em Gamowem i innymi, do najgenialniejszych fizyków teoretycznych Związku Radzieckiego. Jego los był tragiczny: w 1938 roku padł on ofiarą stalinowskiego terroru. Bronstein zajmował się między innymi eksperymentami myślowymi w odniesieniu do pomiaru kwantyzowanego pola grawitacyjnego. Takie eksperymenty myślowe stały się modne w dziedzinie teorii kwantowej po tym, jak Heisenberg, a potem Bohr i Einstein często korzystali z nich w związku z zasadami nieoznaczoności. Zgodnie z tymi zasadami istnieją sprzężone wielkości, takie jak położenie i pęd, których nie można jednocześnie zmierzyć. Pojedyncze wielkości, jak samo położenie lub sam pęd, można natomiast zmierzyć z dowolną dokładnością. Z elektrodynamiki kwantowej wynikały jednak dużo drastyczniejsze ograniczenia. W 1931 roku Landau i pochodzący z Niemiec fizyk Rudolf Peierls (1907–1995) twierdzili, że niektórych wielkości, jak na przykład kwantyzowanego pola elektrycznego, nie można zmierzyć z dowolną dokładnością. Czy teoria taka, mówiąca o tak poważnych ograniczeniach, rzeczywiście ma sens? W 1933 roku Bohr wraz z belgijskim fizykiem Léonem Rosenfeldem (1904–1974) poświęcili temu problemowi obszerną pracę. Obaj doszli do wniosku, że w przypadku elektrodynamiki nie istnieją zasadnicze ograniczenia co do pomiaru pojedynczych wielkości pola, lecz obowiązują tylko te przewidziane przez zasady nieoznaczoności, dotyczące wielkości sprzężonych. (W elektrodynamice pole elektryczne i magnetyczne są wielkościami sprzężonymi podobnie jak położenie i pęd, zatem nie można ich mierzyć jednocześnie z dowolną dokładnością). W eksperymentach myślowych zostało to wyrażone tak, że ładunkiem elektrycznym próbek można było dysponować właściwie w dowolny sposób.

Do tego nawiązał Bronstein. Zastosował on eksperyment myślowy, omawiany przez Bohra i Rosenfelda, w odniesieniu do pomiaru pola grawitacyjnego. W tym wypadku decydującą rolę odgrywa masa, a nie ładunek. Masą nie można jednak dysponować w sposób dowolny. Zbyt duża masa może doprowadzić do tego, że promień Schwarzschilda dla danego ciała przekroczy jego wymiary i ciało stanie się czarną dziurą (jest to pojęcie, którego Bronstein jeszcze nie znał). Bronstein wymagał więc, aby masa próbki nie przekraczała tej granicy. W ramach swojego eksperymentu myślowego uznał on, że pola grawitacyjnego nie można zmierzyć z dowolną dokładnością. Na tej podstawie wywnioskował również, że musi istnieć zasadnicze ograniczenie co do ważności geometrii Riemanna – grawitacja jest przecież niczym innym jak geometrią. Bronstein stwierdził, że być może musimy porzucić nasze tradycyjne wyobrażenia o przestrzeni oraz czasie i zastąpić je bardziej zasadniczymi i zupełnie nieoczekiwanymi pojęciami. Oczywiście w wyniku badań Bronsteina nie dowiedziono, że nie ma sytuacji, w której dokładny pomiar pola grawitacyjnego nie byłby możliwy. Dzięki nim wiemy jednak, że korzystając z tradycyjnych metod, dochodzimy do granicy naszych możliwości. Później przekonamy się, że według aktualnych ujęć na temat grawitacji kwantowej w przestrzeni istnieje najmniejsza długość, poniżej której klasyczna geometria traci swoje znaczenie.

Nawet jeśli wiele argumentów przemawia za kwantyzacją grawitacji, nie oznacza to automatycznie logicznej konsekwencji jej istnienia. W związku z tym nowi autorzy co pewien czas próbują sformułować fundamentalną teorię, w której klasyczne właściwości pola grawitacyjnego pozostawałyby nienaruszone, a jednocześnie wszystkie inne stopnie swobody byłyby opisane przez pola kwantowe. Teorie takie określa się niekiedy mianem „semiklasycznych". Jak dotąd nie odniosły one spektakularnego sukcesu. Najprostszą z nich można nawet obalić za pomocą jednego eksperymentu (który rzeczywiście przeprowadzono w analogicznej formie). Jak brzmi najprostsza teoria semiklasyczna? Równania pola grawitacyjnego autorstwa Einsteina składają się z dwóch części (zob. część 1): jedna dotyczy geometrii, a druga – materii (wszystkie niegrawitacyjne stopnie

swobody). Obie części tych równań należy uznać za klasyczne. Z uwagi jednak na to, że materia spełnia wymogi teorii kwantowej, równania te nie mogą być poprawne w aktualnej formie. Jedna z możliwości, aby uratować z tych równań jak najwięcej, polega na zastąpieniu części dotyczącej materii kwantową wartością średnią z wszystkich możliwych wyników pomiarów, które dopuszcza stan kwantowy. Taka wartość średnia jest ponownie wielkością klasyczną, która z kolei może służyć za podstawę klasycznego pola grawitacyjnego.

W 1981 roku amerykańscy fizycy Don Page i C.D. Geilker opracowali pewien eksperyment pozwalający przetestować tę semiklasyczną teorię. Koncepcja przypomina trochę eksperyment myślowy Schrödingera, którego głównym bohaterem był kot (rys. 5.1). We wnętrzu pojemnika znajduje się materiał promieniotwórczy wraz z dwiema masami połączonymi ze sobą za pomocą sprężyny. Początkowo między masami znajduje się jeszcze sztywne połączenie, ale w wyniku rozpadu promieniotwórczego zostaje ono przerwane i masy mogą kołysać się w dowolnych kierunkach. (Przerwanie sztywnego połączenia oznaczono na rysunku za pomocą krzyżyka). Poza pojemnikiem znajduje się tak zwana waga Cavendisha, za pomocą której można zmierzyć pole grawitacyjne mas, a tym samym również rozkład masy.



5.1. Eksperyment Page'a i Geilkera

Zgodnie z teorią semiklasyczną pole grawitacyjne zależy od wszystkich komponentów funkcji falowej, które opisują działające masy. Wartość średnia ma obejmować wszystkie wyniki pomiarów, a zatem zarówno stan „Rozpad materiału promieniotwórczego", jak i „Brak rozpadu materiału promieniotwórczego". Wskazanie wagi Cavendisha opierałoby się na tej wartości średniej rozkładu masy, która tylko stopniowo zmieniałaby się w czasie. Z kolei zgodnie z teorią grawitacji kwantowej waga ta reagowałaby na poszczególne komponenty, czyli „rozkład" lub „brak rozkładu". Działoby się tak, dopóki nie doszłoby do zmiany liniowego charakteru teorii kwantowej, czyli do czasu obowiązywania zasady superpozycji. (Jest to równoznaczne z przyjęciem hipotezy Everetta, w ramach której kolaps funkcji falowej prowadziłby do problemów z zachowaniem konsystencji). W takim wypadku waga oraz każdy z jej obserwatorów reagowaliby na komponenty „rozkład" i „brak rozkładu". W pewnym momencie doszłoby zatem do wychylenia wskazówki wagi, co odpowiadałoby przerwaniu sztywnego połączenia między masami w wyniku wystąpienia rozpadu promieniotwórczego. Właśnie to zachowanie udało się zaobserwować podczas przeprowadzania eksperymentu!

Wynik ten zapewne nie na wszystkich zrobi wrażenie, ponieważ odpowiada on powszechnemu oczekiwaniu co do rezultatu eksperymentu (wychylenie wskazówki w momencie ruchu mas). W odniesieniu do koncepcji teoretycznych jest on jednak w pewnym stopniu kontrowersyjny, gdyż jego rezultat pokazuje, że doświadczenie jest sprzeczne z pozornie niewinnie wyglądającymi równaniami, takimi jak semiklasyczne równania Einsteina. Eksperyment ten nie dostarcza z pewnością żadnego dowodu na istnienie grawitacji kwantowej, ale pośrednio na to wskazuje.

Jak skwantyzować grawitację?

Dostrzeżenie konieczności istnienia grawitacji kwantowej to jedno, a sformułowanie odpowiedniej teorii – drugie. Największym minusem

w poszukiwaniach takiej teorii jest to, że w dużej mierze brakuje wskazówek dostarczanych przez eksperymenty. Fakt ten skrywa oczywiście niebezpieczeństwo zatracenia się w teoriach, które nie mają żadnego związku z rzeczywistością. Na przykład do sformułowania teorii kwantowej nigdy by nie doszło, gdyby nie powstała potrzeba jej istnienia w celu wyjaśnienia faktów eksperymentalnych (promieniowanie wnękowe, widma). Podczas poszukiwań należy zatem podążać za zdefiniowanymi pojęciami i matematycznymi stanowiskami, nie zapominając jednocześnie o eksperymentach.

Z uwagi na założenie uniwersalności teorii kwantowej efekty działania grawitacji kwantowej powinny być widoczne właściwie w każdej skali. Powyższe ma już swoje odzwierciedlenie w mechanice kwantowej, w ramach której zasada superpozycji nie podlega początkowo żadnym ograniczeniom. Fakt, iż mimo to w ujęciu makroskopowym przyjmuje się zazwyczaj klasyczny obraz świata, związany jest z dekoherencją (zob. część 2): nieuniknione oddziaływanie z otoczeniem nie pozwala na to, aby lokalnie w obserwowanym obiekcie dostrzec kwantowe właściwości interferencyjne. Dotyczy to w szczególności grawitacji, która łączy się z wszystkimi istniejącymi stopniami swobody. W przypadku grawitacji należy więc spodziewać się doskonale ukształtowanego klasycznego zachowania.

Istnieje jednak pewna wyjątkowa skala, w której efekty kwantowe grawitacji powinny być nieuniknione. Jak dochodzi do jej powstania? Ogólna teoria względności Einsteina składa się z dwóch głównych parametrów: prędkości światła c oraz stałej grawitacyjnej G. Gdy przyjrzeć się dokładniej ich jednostkom, można dojść do wniosku, że na tej podstawie nie sposób stworzyć żadnej wielkości z jednostką jednej długości. Ogólna teoria względności sama w sobie nie wyróżnia się zatem żadną skalą długości. Sytuacja ta ulega zmianie, gdy pod uwagę weźmie się teorię kwantową. Za pomocą stałej Plancka h w zestawieniu z c oraz G można stworzyć długość fundamentalną. Z uwagi na to, że Max Planck

jako pierwszy zwrócił na to uwagę, na jego cześć mówi się o „długości Plancka”.

Ktoś, kto dowiaduje się, że Planck wprowadził tę długość już w 1899 roku, czyli na rok przed narodzinami teorii kwantowej, może być nieco zaskoczony. Jak mogło do tego dojść? W tym czasie Planck był bardzo zdeterminowany, aby zrozumieć działanie promieniowania wnękowego (zob. część 2). Przy określonej długości fali widmo tego promieniowania osiąga maksimum, które było doskonale znane z eksperymentów, choć jego ogólna koncepcja pozostawała niezrozumiała. Dopiero wprowadzenie w 1900 roku kwantu działania dostarczyło wyjaśnienia. Jednak już w 1899 roku na podstawie danych eksperymentalnych Planck stwierdził, że musi istnieć nieznana dotąd wielkość o wymiarze fizycznym działania (czyli ilorazu energii i czasu), którą oznaczył wtedy literą *b*. Wielkość ta pokrywa się z wprowadzonym rok później kwantem działania.

Za pomocą *h*, *G* oraz *c* można zatem stworzyć długość fundamentalną, czyli długość Plancka, ale jest ona mikroskopijna i wynosi zaledwie około 10^{-35} metra! Aby wyobrazić sobie taką wielkość, można w myślach nadmuchać atom (około 10^{-10} metra) do wielkości orbity Księżyca (mającej promień około 400 000 kilometrów). W tej skali długość Plancka z zaledwie około 10^{-17} metra odpowiadałaby najmniejszej obecnie skali długości, odczytywanej w największych akceleratorach na świecie. Jeśli więc długość Plancka jest skalą grawitacji kwantowej, to owo porównanie obrazuje, dlaczego teoria ta nie odgrywa jeszcze żadnej roli empirycznej, skoro daleko jej do zbadanych dotąd wymiarów z zakresu fizyki atomowej, jądrowej czy fizyki cząstek elementarnych. Gdyby istniała cząstka o wymiarach Plancka, to wywołane przez nią zakrzywienie przestrzeni nie byłoby pomijalne. Grawitacja *oraz* teoria kwantowa odrywałyby wtedy taką samą rolę.

Oprócz długości Plancka za pomocą trzech powyższych wielkości można wyznaczyć również jednostki czasu i masy (a tym samym energii), które określa się mianem czasu Plancka i masy Plancka. Czas Plancka jest równie mikroskopijny co długość Plancka i wynosi 10^{-44} sekundy. Masa

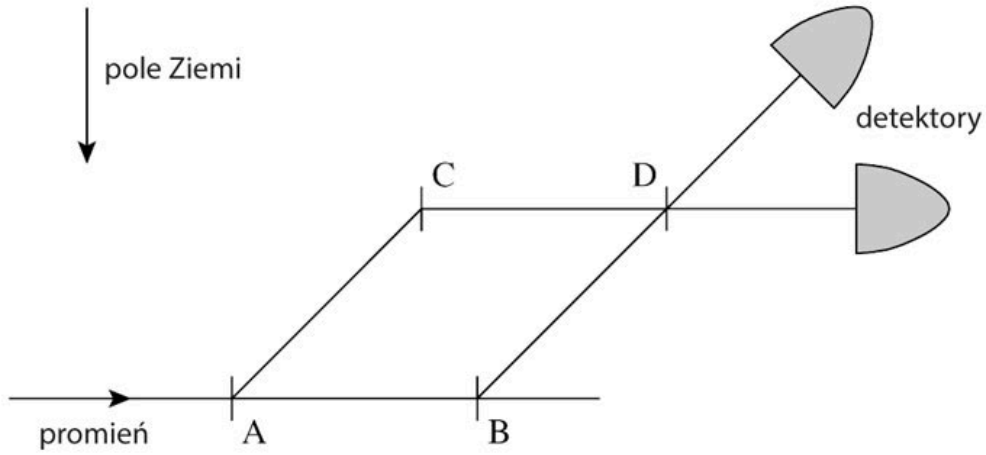
Plancka na pierwszy rzut oka nie wygląda tak egzotycznie i wynosi około jednej stumilionowej (10^{-8}) kilograma, co znacznie przekracza masę na przykład wirusów. W skali masy cząstek jest to faktycznie ogromna masa – masa Plancka to mniej więcej 10^{19} razy więcej niż masa protonu! Odpowiednio duża musi być także energia potrzebna do wytworzenia w akceleratorze cząstki o takiej masie. Do stworzenia takich pokładów energii niezbędny byłby akcelerator o rozmiarach galaktyki. Finansowania takiej budowy nie podjąłby się z pewnością żaden rząd. Nawet akcelerator o rozmiarach Księżyca, o jakim marzył Enrico Fermi, okazałby się niewystarczający.

Z uwagi na to, że prawa natury są takie same w całym wszechświecie, jednostki Plancka są wielkościami uniwersalnymi. Planck był tego świadomy. U kresu swojej pracy w 1899 roku pisał on, że wielkości te zachowują swe naturalne znaczenie, o ile obowiązuje nadal prawo grawitacji, prawo rozchodzenia się światła w próżni oraz dwie podstawowe zasady termodynamiki. Różne formy życia będą mogły je zmierzyć, korzystając z różnych metod pomiaru, bowiem zawsze będą takie same. Wspaniała dalekowzroczność płynie z tych słów!

W poprzedniej części niniejszej książki omówiliśmy kwestię struktur wszechświata. Ich wielkość i masę określa się głównie za pomocą masy protonu, która zajmuje centralne miejsce w dziedzinie fizyki atomowej i jądrowej. Duża rozbieżność między masą protonu a masą Plancka jest odpowiedzialna za to, że ta ostatnia nie odgrywa praktycznie żadnej roli w rozwiązaniach astrofizycznych. Gdyby było inaczej, o grawitacji kwantowej uczono by już na pierwszym semestrze studiów. Jest jednak inaczej, a więc teoria ta jest z jednej strony potrzebna po to, aby przedstawić tak egzotyczne sytuacje jak parowanie czarnych dziur czy wielki wybuch, a z drugiej strony także po to, aby zrozumieć, dlaczego świat jest taki, jaki jest.

W ramach wstępu do grawitacji kwantowej przytoczymy jeszcze krótko, co tak naprawdę wiemy o relacji między grawitacją a teorią kwantową od strony eksperymentalnej. W pierwszej kolejności mamy

zatem mechanikę kwantową atomów i cząstek w polu grawitacyjnym Ziemi. W tym przypadku nie chodzi wprawdzie o kwantyzację pola grawitacyjnego, ale mimo wszystko jest to płaszczyzna, na której przeprowadza się już eksperymenty. W ten sposób można badać zjawiska interferencji zachodzące wtedy, gdy rozszczepi się pojedynczy atom lub neutron na dwie wiązki, które przed ponownym połączeniem w jedną całość wpadają do pola Ziemi na dwóch różnych wysokościach (zob. rys. 5.2). Dzięki temu do wzorca interferencyjnego trafiają szczegóły dotyczące pola grawitacyjnego Ziemi. (Za pomocą tego eksperymentu można nawet dowieść ruchu obrotowego Ziemi, dzięki czemu zbędne staje się przedstawione na rysunku 1.14 wahadło Foucaulta). W ten sposób można jednak przede wszystkim stwierdzić, że zasada równoważności (czyli równość masy grawitacyjnej i bezwładnej) obowiązuje także w ramach kwantowych, co na początku eksperymentu wcale nie było takie oczywiste. Dzięki innym eksperymentom udało się zaobserwować dyskretne stany energetyczne neutronów w polu grawitacyjnym, których istnienie zakładała mechanika kwantowa.



5.2. Zasada działania eksperymentu w zakresie interferencji atomów lub neutronów w polu Ziemi: cząstka napływa z lewej strony i w wyniku działania dzielnika wiązki dochodzi do jej rozszczepienia w punkcie A. Obie wiązki trafiają do pola Ziemi na różnej wysokości, a następnie łączą się ponownie w punkcie D; punkty B i C symbolizują pozostałe dzielniki wiązek. Wzorzec interferencyjny zostaje potwierdzony w detektorach.

Kolejna płaszczyzna nie doczekała się wprawdzie żadnych eksperymentów, ale dostępne są za to konkretne prognozy. Chodzi tu o omawiany przez nas w części 2 tej książki efekt Hawkinga dla czarnych dziur oraz jego odpowiednik dla płaskiej przestrzeni – również wspomniany tam efekt Unruha, a przedmiotem prognoz jest płaszczyzna pól kwantowych w ściśle ustalonej i w znacznym stopniu zakrzywionej czasoprzestrzeni. Dopiero po rezygnacji z tej czasoprzestrzeni jako tła i poddaniu jej kwantyzacji można osiągnąć pełną płaszczyznę grawitacji kwantowej.

W odniesieniu do tego, co napisano powyżej, brak tła jest tu jednak czynnikiem decydującym. W teoriach niegrawitacyjnych pola opierają się zawsze na niedynamicznej czasoprzestrzeni. W ogólnej teorii względności Einsteina geometria czasoprzestrzeni jest polem grawitacyjnym, co czyni ją dynamiczną. Już sam Einstein łamał sobie nad tym głowę. Opracowanie jego teorii opóźniło się o dwa lata, ponieważ pierwotnie wierzył, że punkty

w czasoprzestrzeni nie wiążą się z polem grawitacyjnym. Dla matematyka ma to może jakiś sens, ale dla fizyka już nie. Pole grawitacyjne oparte na geometrii wszechświata to już wszystko; nie istnieją inne wolne punkty. Podczas transformacji współrzędnych należy odwzorować w równej mierze wszystkie punkty, jak i metryki oraz pola materii – nie można myśleć o polach w oderwaniu od punktów. Dopiero po zidentyfikowaniu swojego pierwotnego błędu Einstein w 1915 roku uzyskał prawidłowe równania pola grawitacyjnego. Stwierdzenie to jest bardzo istotne w kontekście grawitacji kwantowej, gdyż wymaga nowych metod kwantyzacji, a tym samym również – kwantyzacji czasoprzestrzeni.

Jak zatem skwantyzować grawitację? Istnieją dwa główne nurty. W pierwszej kolejności można pomyśleć o dokonaniu kwantyzacji ogólnej teorii względności Einsteina bez konieczności zapewnienia unifikacji z innymi oddziaływaniami. Jest to kierunek, który będziemy omawiać w trzech kolejnych rozdziałach. Niektórzy fizycy przypuszczają jednak, że osiągnięcie celu jest możliwe tylko wtedy, gdy od początku założy się zunifikowaną teorię kwantową wszystkich oddziaływań, w ramach której nie doszło jeszcze do odseparowania grawitacji od innych oddziaływań. Najambitniejszym i do tej pory jedynym programem realizującym tę myśl jest teoria strun, którą także później poznamy. Wszystkie ujęcia mają swoje mocne i słabe strony. Przekonamy się, co mają one do powiedzenia w odniesieniu do najbardziej palących i nierozwiązanych problemów naszego wszechświata.

Grawitacje, nieskończoności, całki po trajektoriach

Lata dwudzieste XX wieku obfitowały w sukcesy z dziedziny fizyki. W 1927 roku założenia mechaniki kwantowej były właściwie ukończone, a jej pionierzy koncentrowali się nie tylko na ekscytujących możliwościach z zakresu atomu i cząsteczki, lecz również na wyodrębnieniu kwestii pola elektromagnetycznego, co skutkowało potem powstaniem elektrodynamiki kwantowej (zob. część 2). Bardzo istotna w tym kontekście okazała się praca Heisenberga i Pauliego na temat dynamiki pól kwantowych,

opublikowana w 1929 roku w czasopiśmie „Zeitschrift für Physik” pod tytułem *Zur Quantendynamik der Wellenfelder*. Po przedstawieniu niezbędnych dla elektrodynamiki norm autorzy piszą także, że kwantyzację pola grawitacyjnego, która ze względów fizycznych wydaje się konieczna, można przeprowadzić bez doświadczania nowych trudności zgodnie z analogicznymi normami.

W Zurychu Pauli poprosił swojego asystenta Léona Rosenfelda, aby ten zajął się sprawą kwantyzacji pola grawitacyjnego. Przy tej okazji miała zostać rozwiązana zagadka, która przyprawiała Heisenberga i Pauliego o ból głowy, a dotyczyła wyniku obliczeń procesów elektromagnetycznych. Wynikiem tym była niekiedy, zupełnie nieoczekiwanie, nieskończoność, a więc rezultat kompletnie pozbawiony sensu, który nic nikomu nie dawał. Heisenberg podejrzewał, że winę za to może ponosić elektron jako cząstka punktowa. Pauli polecił zatem Rosenfeldowi, aby ten obserwował bezpośrednio grawitacyjną energię światła, to jest przypadek pozbawiony materii. Zgodnie ze wzorem $E = mc^2$ Einsteina energia światła jest równa masie, która tworzy pole grawitacyjne. W ramach swoich rozważań Rosenfeld stwierdził jednak, że wartość energii powstałego pola również jest nieskończona. W związku z tym elektron nie może być powodem otrzymania tego nielogicznego wyniku. Przyczyny tego stanu rzeczy muszą być zatem dużo bardziej złożone. W rzeczywistości chodziło tu o ogólny problem kwantowych teorii pól, którym zajmowaliśmy się już w części 2 tej książki i do którego jeszcze wrócimy.

W części 2 poznaliśmy foton jako kwant pola elektromagnetycznego. Posiada on spin = 1 (jest on więc bozonem) i zerową masę spoczynkową oraz jest nośnikiem oddziaływań między ładunkami elektrycznymi. Czy istnieje analogiczny kwant dla pola grawitacyjnego – *grawiton*? Na początku lat trzydziestych XX wieku fizycy, na przykład Bohr i Fermi, przez pewien czas rozważali zaproponowane przez Pauliego neutrino jako kwant pola grawitacyjnego. Koncepcja ta nie była jednak słuszna. Neutrino jest fermionem, a zatem posiada spin = ½. Tymczasem grawiton,

analogicznie do fotonu, jako nośnik oddziaływań dalekiego zasięgu, powinien być bozonem. Wiedział o tym również Bronstein, który we wspomnianej już wcześniej pracy, opublikowanej w 1936 roku w czasopiśmie fizycznym wydawanym w Związku Radzieckim, opisał kwantyzację słabych fal grawitacyjnych. Bronstein pokazał, że teoria grawitacji Newtona jest przybliżeniem oddziaływań między masami, których nośnikiem są grawitony, analogicznie do aproksymacyjnego obowiązywania prawa Coulomba w wyniku wymiany fotonów między ładunkami.

W związku z tym można więc dostrzec bliską analogię między grawitacją a elektromagnetyzmem. Istnieje jednak również jedna podstawowa różnica: ze względu na działanie grawitacji masy zawsze się przyciągają, natomiast ładunki mogą zarówno się przyciągać, jak i odpychać. W rzeczywistości istnieją wyłącznie masy dodatnie, podczas gdy ładunki mogą być zarówno dodatnie, jak i ujemne. W odróżnieniu od mas ładunki jednego znaku nie przyciągają się, lecz odpychają. Ową różnicę można uzasadnić tym, że grawiton w odróżnieniu od fotonu posiada spin = 2. Kwestię tę w ważnej publikacji poruszyli Pauli i Fierz w 1939 roku: równania, które można znaleźć dla cząstek bezmasowych o spinie = 2, to po prostu równania Einsteina w przybliżeniu liniowym. Heisenberg był tym faktem zachwycony. Tuż przed wybuchem drugiej wojny światowej z miejscowości wypoczynkowej nad Walchensee, gdzie spędzał urlop, napisał on do Pauliego, że kwanty grawitacyjne o spinie = 2 sprawiły mu frajdę. Potem wszyscy mieli już inne zmartwienia...

Co się stało z tymi dziwnymi nieskończonościami (dywergencjami), na które natknęli się Heisenberg, Pauli i Rosenfeld? Łączą się one bezpośrednio ze wspomnianą już w części 2 właściwością pól kwantowych, dotyczącą posiadania nieskończenie wielu stopni swobody, ponieważ jedno pole w każdym punkcie czasoprzestrzeni ma odrębną wartość, a punktów jest nieskończenie wiele. Teoria mówi o dowolnie krótkich odstępach, a tym samym – zgodnie z zasadą nieoznaczoności – również o dowolnie wysokich pędach i energiach. To właśnie te dowolnie

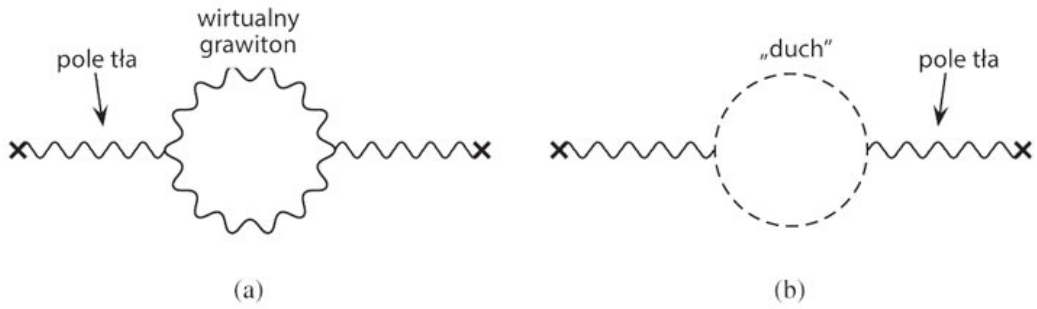
wysokie energie są odpowiedzialne za występowanie dywergencji. Mimo że nieskończoności są obecne w praktycznie każdej kwantowej teorii pola, istnieją istotne różnice: dokonano podziału na teorie *renormalizowane* i *nierenormalizowane*. W przypadku teorii renormalizowanych nieskończoności znajdują się w pewnym sensie pod kontrolą. Typowym przykładem jest elektrodynamika kwantowa. W jej przypadku dzięki pomysłowemu zabiegowi matematycznemu wszystkie dywergencje można zamknąć w redefinicji („renormalizacji") dwóch parametrów: masy i ładunku elektronu. Parametrów tych nie można rzecz jasna obliczyć, a więc trzeba je określić eksperymentalnie. Ich celem jest wyrażenie nieskończoności. Teorie silnego i elektrosłabego oddziaływania, czyli model standardowy fizyki cząstek elementarnych, są także renormalizowane, co jest jedną z przyczyn sukcesu, jaki odniosły.

W przypadku teorii nierenormalizowanych sytuacja jest dużo mniej komfortowa. Istnieje tak wiele rodzajów nieskończoności, że potrzebna byłaby praktycznie nieskończona ilość parametrów, aby je wchłonąć, a zatem nieskończenie wiele wielkości, które trzeba obliczyć. Wszystko to sprawia, że teoria wydaje się mało wiarygodna. Teorię grawitacji kwantowej, stworzoną w ramach analogii do elektrodynamiki kwantowej, uznaje się niestety za teorię nierenormalizowaną. Czy taka teoria jest coś warta?

Wszystkie wymienione powyżej nieskończoności występują w tak zwanej „teorii perturbacji" – jest to przypadek graniczny, w którym skwantyzowane pola tworzą niewielkie perturbacje wokół ściśle określonego tła. Większość kwantowych teorii pól jest zdana na ten przypadek graniczny, ponieważ są one zbyt skomplikowane, aby pozwolić na badanie pól o dowolnej wielkości. Zgodnie z tym, o czym pisaliśmy powyżej, założenie istnienia tła jest niezgodne z duchem teorii względności – przecież również czasoprzestrzeń podlega kwantyzacji. W związku z tym przypuszczano, że teoria perturbacji w przypadku grawitacji kwantowej nie ma sensu i dlatego trzeba opracować koncepcję, która nie bierze pod uwagę tła. Koncepcje te omówimy później bardziej

szczegółowo. Mimo to teorii nierenormalizowanych nie postrzega się obecnie tak negatywnie jak kiedyś. Powodem tego stanu rzeczy jest to, że czasami także na podstawie takich teorii mogą powstać konkretne prognozy – pod warunkiem założenia odpowiednio niewielkich energii. W przypadku grawitacji kwantowej udało się na przykład obliczyć drobne korekty potencjału newtonowskiego, które są proporcjonalne do kwantu działania *h*. Wprawdzie odchylenia są zbyt małe, aby w przewidywalnym czasie móc zaobserwować je w laboratorium (potrzebne byłyby pomiary odległości w skali Plancka), jednak są to konkretne efekty działania grawitacji kwantowej, które można obliczyć i przynajmniej *w teorii* zmierzyć.

Do obliczenia takich efektów wykorzystuje się wspomniane w części 2 tej książki diagramy Feynmana, które pozwalają przełożyć skomplikowane procesy z dziedziny teorii kwantowej na intuicyjny język znaków. Rysunek 5.3 to przykład dwóch diagramów, które należy uwzględnić w rozważaniach nad teorią perturbacji w grawitacji kwantowej. Z lewej strony można dostrzec wirtualny grawiton oddziałujący z obecnym w teorii perturbacji tłem. Czym jest jednak wirtualna cząstka widoczna w pętli po prawej stronie? W rzeczywistości nie jest to „cząstka”, lecz tylko wyrażenie określonych termów, które trzeba uwzględnić podczas obliczeń. W związku z tym cząstkę tę nazywa się czasami także „duchem”. To zdecydowana korzyść płynąca z języka Feynmana: wielkości z równań matematycznych, których nie da się pokazać, można przedstawić jako wirtualne cząstki – cząstki, które w rzeczywistości nie istnieją.



5.3. Dwa diagramy Feynmana, które są niezastąpione podczas obliczeń procesów z dziedziny grawitacji kwantowej

Richard Feynman nie tylko był pomysłodawcą nazwanych jego nazwiskiem diagramów, lecz również przyczynił się do przeformułowania teorii kwantowej. Pokazał on bowiem, że funkcję falową z zakresu mechaniki kwantowej można obliczyć nie tylko za pomocą równania Schrödingera. W tym celu należy zsumować całą drogę (w tym kontekście nazywaną *trajektorią*), jaką mogła przebyć cząstka między dwoma wyznaczonymi miejscami i czasami – oczywiście może to być więcej niż jedna cząstka lub jedno pole. W ramach tej sumy każda trajektoria otrzymywała odpowiednią wagę wyrażoną za pomocą określonego czynnika matematycznego. Sumę tę nazywa się też „całką po trajektoriach". Trajektorie są oczywiście fikcyjne – w teorii kwantowej nie ma przecież żadnych torów. Można jednak *założyć* istnienie na przykład elektronu, który jednocześnie przebywa wszystkie odcinki, aby w ten sposób wyjaśnić wzorzec interferencyjny zaobserwowany w eksperymencie (i przedstawiony na rysunku 2.2).

Z całek po trajektoriach można również skorzystać w ramach grawitacji kwantowej. W tym przypadku zamiast całej możliwej drogi cząstki należy zsumować całą możliwą czasoprzestrzeń wraz z jej metrykami – doprawdy niełatwe przedsięwzięcie. W ten sposób można jednak, przynajmniej w założeniu, wykroczyć poza teorię perturbacji, otrzymując koncepcję bez tła.

Jak obliczyć całkę po trajektoriach w ramach grawitacji kwantowej? Bardzo obiecującą ewentualnością jest podział czasoprzestrzeni, po których trzeba sumować, na dyskretne części i późniejsze ich zsumowanie. W jednym z ujęć części te tworzą sympleksy równoboczne. Czym są sympleksy? Sympleks jest uogólnieniem trójkąta na więcej wymiarów. W naszym przypadku powierzchnia sympleksu składa się z trójkątów równobocznych. W dwóch wymiarach jest to trójkąt, a w trzech – czworościan, czyli trójkątna piramida złożona z czterech trójkątów. Czterowymiarowy sympleks lub krótko 4-sympleks uzyskuje się poprzez dodanie do czterościanu punktu, który znajduje się w czwartym wymiarze i łączy się z czterema punktami czworościanu, dając w efekcie wielokomórkę o pięciu wierzchołkach, dziesięciu ścianach trójkąta i pięciu wielościanach. Rysunek 5.4 przedstawia 4-sympleks foremny, a dokładniej jego rzut na płaszczyznę papieru. Analiza całki po trajektoriach w grawitacji kwantowej sprowadza się zatem do problemu kombinatorycznego, czyli do poprawnego zsumowania wszystkich właściwych 4-sympleksów.

5.4. Rzut 4-sympleksu foremnego na płaszczyznę papieru

Nie jest to z pewnością rzecz łatwa, mimo to przy wykorzystaniu komputera można spróbować przeprowadzić odpowiednie obliczenia. Ostatecznie najkorzystniejsze byłoby oczywiście stopniowe zmniejszanie długości trójkątów bocznych i rozpatrzenie przypadku granicznego, w którym osiąga ona wartość niemal zerową – rzeczywiste czasoprzestrzenie, po których ma odbywać się sumowanie, nie mają

przecież żadnych wierzchołków ani krawędzi. Jak dotąd nie udało się sprawdzić przypadku granicznego. W przeciwnym razie za pomocą tej metody doprowadzono by do rozwiązania problemu grawitacji kwantowej. Mimo to dostępne są już interesujące wyniki. Zgodnie z nimi czasoprzestrzeń w najmniejszych skalach – zbliżonych do skali Plancka – jest raczej dwu-, a nie czterowymiarowa. Nasza intuicja zawodzi przy tak niewielkich odległościach.

Kanoniczna grawitacja kwantowa i problem czasu

Głównym problemem w przypadku kwantyzacji grawitacji jest znalezienie koncepcji, która nie wymaga tła, czyli nie mieści się w ramach ściśle ustalonej czasoprzestrzeni. Ujęcie, które ze względu na swoje główne założenia zdaje się najbardziej zbliżone do tych oczekiwań, określa się mianem kanonicznej grawitacji kwantowej. Co należy przez to rozumieć? Słowo „kanon" pochodzi z języka greckiego. Pierwotnie oznaczało ono pręt służący do wykonywania pomiarów. Później znaczyło ono tyle co zbiór reguł, odpowiadający normie. W teorii kwantowej mianem tym określa się popularną koncepcję opierającą się na położeniach i pędach oraz stanach kwantowych (funkcjach falowych), które są zależne od położenia *lub* pędu (zależność od obu tych czynników jednocześnie jest ze względu na zasadę nieoznaczoności niemożliwa). Istnieją też koncepcje kanoniczne dla pól, przy czym jednak zamiast położenia pojawia się w nich wartość pola.

Warunkiem stosowania kanonicznej grawitacji kwantowej jest uprzednie opracowanie koncepcji w ramach ogólnej teorii względności, która posługiwałaby się pojęciami kanonicznymi. Nie jest to wcale takie proste. Teoria Einsteina opiera się w dużej mierze na pojęciu czasoprzestrzeni jako całości, ponieważ jej geometria opisuje pole grawitacyjne. Co może być zatem odpowiednikiem uogólnionych położeń i pędów? W mechanice klasycznej to tor cząstki tworzy porządek miejsc w czasie. Analogicznie jest w przypadku czasoprzestrzeni – ją również można rozumieć jako porządek trójwymiarowych przestrzeni. Tor cząstki

jest więc odpowiednikiem czasoprzestrzeni, a miejsce cząstki – odpowiednikiem przestrzeni. W kanonicznej koncepcji teorii względności to przestrzeń przejmuje zatem rolę zmiennej, od której zależy funkcja falowa, a dokładniej: zmienna ta jest metryką trójwymiarowej przestrzeni (w skrócie: 3-metryką). Co jest odpowiednikiem pędu? Okazuje się, że dostarcza go wielkość geometryczna określana mianem „krzywizny zewnętrznej" (porównaj z częścią 1). Wielkość ta informuje o tym, w jaki sposób trójwymiarowa przestrzeń jest osadzona w czwartym wymiarze.

W przypadku powierzchni, które są osadzone w przestrzeni, odpowiednikiem sprzężonych wielkości 3-metryki i krzywizny zewnętrznej są wprowadzone przez Gaussa wielkości krzywizny wewnętrznej i zewnętrznej. Ze względu na osadzenie powierzchni w przestrzeni obie te wielkości nie mogą być od siebie niezależne, ponieważ nie każdą dowolną powierzchnię można umieścić w przestrzeni. Związek między krzywizną wewnętrzną i zewnętrzną Gauss określił mianem *theorema egregium* (z pojęciem tym spotkaliśmy się już w części 1). W odniesieniu do teorii względności, w której trójwymiarowa przestrzeń jest wpisana w czterowymiarową czasoprzestrzeń, odpowiednią wersję tego twierdzenia nazywa się „warunkiem koniecznym Hamiltona".

William Rowan Hamilton (1805–1865) był irlandzkim matematykiem i fizykiem o wszechstronnych zainteresowaniach. Jednym z jego największych osiągnięć było przeformułowanie mechaniki klasycznej, dla której uszył zupełnie nowe i dłuższe matematyczne szaty. Zdumiewa fakt, że matematyka, na której opiera się mechanika Hamiltona, jest taka sama jak w przypadku optyki geometrycznej. Optyka geometryczna jest przypadkiem granicznym optyki falowej, w którym wystarczy wykorzystanie promieni świetlnych. Przypadek ten zachodzi wtedy, gdy długość fali jest znacznie mniejsza od wymiarów aparatury, z którą światło oddziałuje, a zatem właściwości fali są niezauważalne. Ta analogia między optyką a mechaniką była jednym z najważniejszych motywów przewodnich Schrödingera podczas formułowania jego równań mechaniki kwantowej. Gdy optyka geometryczna jest zaledwie przypadkiem

granicznym optyki falowej, to czy w takim razie mechanika klasyczna nie powinna być również przypadkiem granicznym bardziej ogólnej teorii – mechaniki falowej? Myśl ta zaprowadziła Schrödingera do jego funkcji falowej, co przyczyniło się do powstania głównego równania w mechanice kwantowej – równania Schrödingera.

Szczegółowa analiza pokazuje, że sytuacja jest skomplikowana, ale w dużym uogólnieniu nie różni się ona tak znacznie od teorii względności, poza jednym ważnym wyjątkiem – w mechanice i mechanice kwantowej istnieje czas absolutny *t*, a w teorii Einsteina nie. Z tego powodu w funkcji falowej Schrödingera oprócz położenia pojawia się też dodatkowo *t*. W grawitacji kwantowej *t* nie występuje, ponieważ czas absolutny nie istnieje, zatem oczekujemy wyłącznie trójwymiarowej metryki jako argumentu funkcji falowej, będącego zmienną, która pełni funkcję położenia. Powyższe wynika z formalizmu kanonicznego. Parametr czasu *t* znika!

Jak należy rozumieć brak czasu? Rzecz jasna, wynika to z formalizmu, ale czy istnieje również jakieś bardziej obrazowe wyjaśnienie? Owszem, istnieje. Ze względu na zasadę nieoznaczoności mechanika kwantowa nie zna torów cząstek, a tylko funkcje falowe. W teorii względności odpowiednikiem toru cząstki jest cała czasoprzestrzeń rozumiana jako porządek przestrzeni, a więc czasoprzestrzeń ginie w grawitacji kwantowej podobnie jak tor cząstek – w mechanice kwantowej. Z uwagi na to, że nie istnieje także czas absolutny *t*, pozostaje wyłącznie przestrzeń.

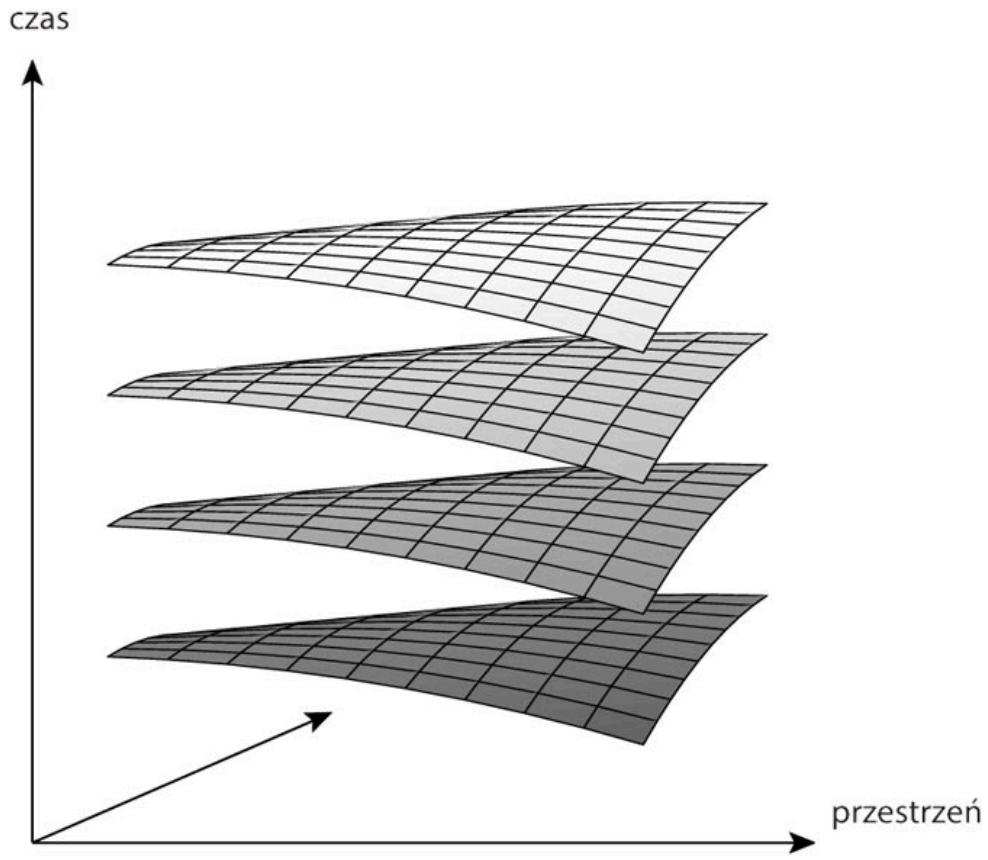
Do tej argumentacji nie są zatem potrzebne szczegóły teorii Einsteina. Każda inna klasyczna teoria grawitacji (jeśli taka w ogóle kogoś interesuje), która nie uwzględnia czasu absolutnego, mówi o bezczasowej grawitacji kwantowej.

Teraz powrócimy na chwilę do kanonicznej koncepcji ogólnej teorii względności. Zgodnie z tym, co podkreślaliśmy już wcześniej, centralną rolę w tej koncepcji odgrywa uogólnienie twierdzenia Gaussa w formie warunku koniecznego Hamiltona. Wyraża on, iż trójwymiarowa przestrzeń może zostać osadzona w czterowymiarowej czasoprzestrzeni.

W rzeczywistości istnieje dość ścisłe powiązanie między trzema i czteroma wymiarami, ponieważ można dowieść, że jeżeli warunek konieczny Hamiltona obowiązuje w przypadku każdej z przestrzeni w czterowymiarowej czasoprzestrzeni, to czasoprzestrzeń musi automatycznie spełniać wymogi równań Einsteina. (Sam warunek konieczny jest już jednym z równań Einsteina). I odwrotnie: jeśli równania Einsteina obowiązują, a warunek konieczny Hamiltona jest spełniony dla jednej przestrzeni, to w przypadku każdej z pozostałych przestrzeni musi być tak samo. W uproszczeniu można zatem powiedzieć, że z „prawa chwili" (warunku koniecznego przestrzeni) wynikają prawa dynamiczne (równania Einsteina dla czasoprzestrzeni). W pewnym sensie warunek konieczny jest więc wystarczający. (W rzeczywistości istnieją jeszcze trzy inne warunki konieczne, których nie będziemy tutaj omawiać). Na rysunku 5.5 dość przekrojowo pokazano, w jaki sposób można rozwarstwić klasyczną czasoprzestrzeń, tworząc szereg trójwymiarowych przestrzeni. Trzeba z góry podkreślić, że wygląd przestrzeni jest dużo bardziej złożony, niż wynika to z naszego rysunku. Rozwarstwienie czasoprzestrzeni na przestrzenie jest w dużej mierze dowolne, co drastycznie różni się od wyobrażenia Newtona (rys. 1.1), które zakładało równy podział na płaskie przestrzenie.

W grawitacji kwantowej czasoprzestrzeń i czas znikają – istnieje tylko przestrzeń. Wszystko, co pozostaje, jest warunkiem koniecznym. Jak należy go traktować w grawitacji kwantowej? W tej kwestii decydujący impuls dał Paul Dirac. Zaproponował on, aby przełożyć warunek konieczny bezpośrednio na warunek konieczny dla funkcji falowej kanonicznej grawitacji kwantowej. Rezultat tej propozycji wygląda na pierwszy rzut oka jak równanie Schrödingera bez czasu, co nie zaskakuje, ponieważ czas *t* w grawitacji kwantowej nie występuje. Funkcja falowa opisuje zatem falę statyczną, która po prostu *jest*, ale w przeciwieństwie do funkcji Schrödingera nie rozwija się w czasie. W jakiej przestrzeni zdefiniowana jest ta funkcja falowa? (Przestrzeń taką nazywa się przecież również przestrzenią konfiguracyjną; zob. część 2). Zgodnie z tym, co widzieliśmy w przypadku koncepcji kanonicznej, odpowiednikiem

zwykłego położenia jest metryka w przestrzeni trójwymiarowej, czyli 3-metryka. W razie potrzeby dochodzą do tego także niegrawitacyjne stopnie swobody, tak jak na przykład pole elektromagnetyczne. Funkcja falowa jest więc określona w przestrzeni wszystkich 3-metryk. Tę skomplikowaną przestrzeń nazywa się „superprzestrzenią". Ta budząca respekt nazwa ma związek z tym, że superprzestrzeń posiada nieskończenie wiele wymiarów. Każdy punkt w superprzestrzeni zawiera 3-metrykę, która jest określona w każdym punkcie zwykłej przestrzeni, a zatem przedstawia nieskończenie wiele liczb. Jak można się domyślić, koncepcja ta nie należy do łatwych.



5.5. Schemat podziału czasoprzestrzeni na szereg przestrzeni

To równanie Schrödingera bez czasu nazywane jest też równaniem Wheelera-DeWitta, na cześć fizyków, którzy je opracowali w latach 1967–

1968 i jako pierwsi przeanalizowali jego konsekwencje. Z Johnem Wheelerem mieliśmy styczność już wcześniej, między innymi w związku z właściwościami czarnych dziur. Bryce DeWitt (1923–2004) napisał ważne prace na temat omawianej w poprzednim rozdziale teorii perturbacji.

Dokładniejsza analiza pokazuje istotną różnicę między równaniem Wheelera-DeWitta a równaniem Schrödingera. W przeciwieństwie do niego ma ono formę równania drgań, takiego jak na przykład w przypadku drgającej struny (choć o dużo wyższym stopniu skomplikowania). Równanie dla drgającej struny zawiera jednak zwykły czas *t*, względem którego struna wykonuje ruch w górę i w dół. Czyż nie stwierdziliśmy przed chwilą, że równanie Wheelera-DeWitta nie zawiera czasu? Zgadza się, o ile mamy na myśli czas zewnętrzny *t*, którego rzeczywiście nie ma. Ze względu na swoją formę (drgania) równanie to wyróżnia jednak wewnętrzną zmienną, która zachowuje się jak zmienna czasu. Jest to część 3-metryki, a mianowicie (w dużym uogólnieniu) ta część, na podstawie której można obliczyć objętość trójwymiarowej przestrzeni. W pewnym sensie zatem wielkość wszechświata dostarcza wymiar czasu – jest to *czas wewnętrzny*, który w odróżnieniu od *t* Newtona nie znajduje się poza wszechświatem, lecz składa się z jego stopni swobody, a w szczególności z jego geometrii.

Koncepcję kanonicznej grawitacji kwantowej, w której funkcja falowa jest zależna od 3-metryki, nazywa się geometrodynamiką kwantową. W następnym rozdziale poznamy alternatywną wersję koncepcji kanonicznej, która posługuje się tak zwanymi pętlami. Równania Wheelera-DeWitta w geometrodynamice nie można rozwiązać w przypadku ogólnym. Z tego powodu poszukuje się alternatyw. W prostszych sytuacjach, jak w przypadku typowego dla kosmologii ograniczenia do przestrzeni jednorodnych i izotropowych, równanie to będzie miało rozwiązanie – opisuje ono wtedy bezpośrednio kosmos kwantowy na poziomie podstawowym (zob. część 6). Za pomocą dokładnego równania Wheelera-DeWitta można mimo to zrozumieć,

w jakich okolicznościach możliwe jest odnalezienie przybliżonego parametru czasu *t*.

Do poszukiwań straconego czasu posłuży nam znaleziona przez Schrödingera analogia między optyką falową a mechaniką kwantową – tym razem jednak w odwrotnym kierunku. Przybliżenie, z którym ponownie pojawia się dawny czas *t*, odpowiada przybliżeniu promieni świetlnych w bardziej podstawowej optyce falowej. Odpowiednikiem promieni świetlnych w mechanice kwantowej są tory cząstek, a w grawitacji kwantowej – czasoprzestrzenie (a tym samym także czas). W optyce falowej można opisać promienie świetlne, gdy długość fali obserwowanego światła jest dużo mniejsza od wymiarów (odległości między szczelinami, średnice otworów itd.) aparatury służącej do obserwacji światła. W grawitacji kwantowej odpowiada to przypadkowi granicznemu, w którym odpowiednie różnice długości są dużo mniejsze od długości Plancka (albo odpowiednie masy są dużo większe od masy Plancka). Powstały parametr *t* nazywa się również „czasem semiklasycznym".

Prostą analogią do powstawania czasu jest czytanie książki. Książka jako całość, taka jaką trzymamy w ręce, jest odpowiednikiem bezczasowej funkcji falowej, ponieważ jej zawartość jest kompletna, a przez to niezależna od czasu. Podczas czytania książki mimowolnie odnosimy wrażenie, że akcja się rozwija – jest to szczególnie widoczne w przypadku kryminałów (chociaż mam nadzieję, że moja książka także buduje napięcie). Owo wrażenie upływającego czasu bierze się stąd, że cały system składa się z dwóch części: książki i czytelnika. To powiązanie między świadomością czytelnika a słowami i zdaniami sprawia, że odnosi on wrażenie upływającego czasu. Podobnie jak większość porównań, kuleje i to, niemniej jednak dotyka ono bardzo ważnej sprawy. Są sytuacje, w których grawitacja kwantowa również pozwala na podział na dwie części: makroskopowe pole grawitacyjne (w pierwszej kolejności wszechświata jako całości) i niegrawitacyjne stopnie swobody. Pierwsza część wyznacza w pewnym sensie granicę czasu dla rozwoju drugiej,

a zatem pierwsza zajmuje rolę książki, a druga – czytelnika. Zamiast pytać o stan systemu w jakimś określonym czasie, można zapytać o jego stan, gdy wszechświat miał określoną wielkość. Wszechświat samodzielnie definiuje czas.

Od dawna znany parametr czasu można jednak odnaleźć zaledwie w przypadku określonych funkcji falowych. Zasada superpozycji pozwala na znacznie więcej stanów kwantowych – stanów, które opisują niestandardowe sytuacje, jak w przypadku jednoczesnego istnienia żywego i martwego kota. W kontekście skwantyzowanego pola grawitacyjnego odpowiada to na przykład superpozycji pola grawitacyjnego masy, która jednocześnie znajduje się w dwóch różnych miejscach. Sytuacja ta jest tak samo mało realna jak superpozycja kota. Wiąże się to z tym, że podobnie jak w przypadku kota zaczyna działać dekoherencja (zob. części 2 oraz 6). Grawitacja łączy się tak samo ze wszystkimi formami energii, przez co w otoczeniu istnieje wiele niezaobserwowanych stopni swobody, które mogą przenosić informację o superpozycji. W najprostszym przypadku są to inne cząstki, które rozpoznają różne części pola grawitacyjnego znajdujące się w superpozycji. Zatem również w ramach grawitacji kwantowej możemy zrozumieć, dlaczego w standardowym przypadku obserwujemy klasyczne pole grawitacyjne. Rzecz jasna, występują także sytuacje, w których nie działa dekoherencja i ujawnia się kwantowa natura grawitacji. Są to oczywiście najbardziej interesujące sytuacje, ponieważ pokazują prawdziwy, niczym niezawoalowany kosmos kwantowy. Do tej kwestii wrócimy jeszcze później.

Grawitacja kwantowa z pętlami

Wspomniana w poprzednim rozdziale geometrodynamika kwantowa pozwala nam zrozumieć istotę grawitacji kwantowej. Dotyczy to nie tylko pojęcia czasu – zanik czasu zewnętrznego na poziomie podstawowym i jego ponowne pojawienie się w określonych okolicznościach w formie przybliżenia. Jednak występują również poważne problemy. Zajmujące centralną pozycję równanie Wheelera-DeWitta nie jest, dopóki nie

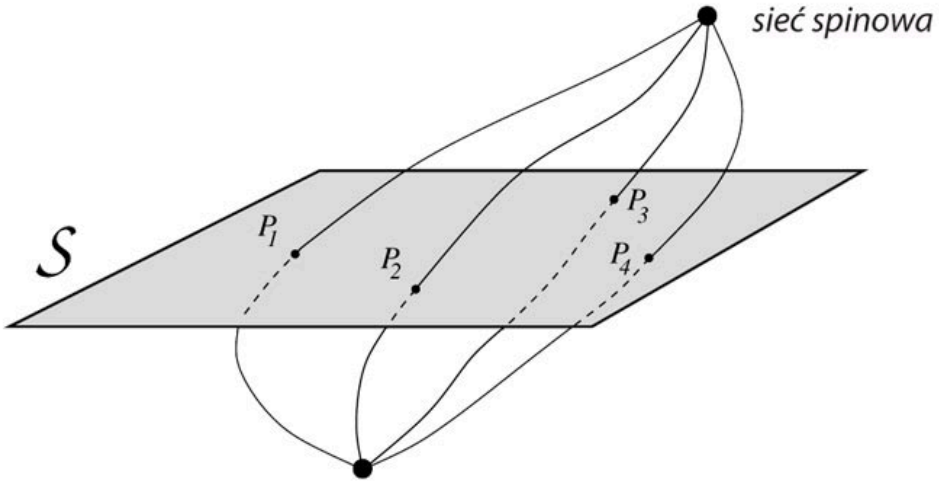
przyjmie się przybliżeń, dobrze zdefiniowane pod względem matematycznym. Wprawdzie można rozpatrywać proste modele (na przykład kosmologii kwantowej, którą omówimy dokładniej w kolejnym rozdziale), w których problemy te nie występują, nie zmienia to jednak faktu, że pełne równanie nie zostanie poddane dokładnej analizie. Winę za to ponosi ponownie zachowanie przy niewielkich odległościach – pojawiają się nieokreślone wyrażenia matematyczne, które uniemożliwiają rozwiązania równania. Nie dowodzi to oczywiście, że rozwiązanie w ogóle nie jest możliwe, ale jak dotąd nie opracowano żadnego wyjścia z tej sytuacji.

W związku z powyższym poszukiwano pomysłu, który wykorzystywałby zalety wynikające z formalizmu kanonicznego, a jednocześnie prowadziłby do rozwiązania problemów matematycznych. W 1986 roku pochodzący z Indii, a pracujący w Stanach Zjednoczonych fizyk teoretyczny Abhay Ashtekar zaproponował, aby w tym celu wprowadzić nowe zmienne. Co to oznacza? Zmienne kanoniczne geometrodynamiki (uogólnione położenia i pędy) stanowią metrykę i krzywiznę zewnętrzną przestrzeni trójwymiarowej. Za pomocą skomplikowanego przekształcenia Ashtekar stworzył z tych zmiennych nowe zmienne, które są analogiczne do wielkości wykorzystywanych w elektrodynamice i fizyce cząstek elementarnych – są one podobne do wspomnianych w części 2 pól cechowania.

Zmienne Ashtekara przeszły kilka zmian od czasu ich wprowadzenia. Obecnie mówi się o „zmiennych pętlowych", podczas gdy odpowiedni formalizm nosi nazwę „pętlowej grawitacji kwantowej". Jak sama nazwa wskazuje, uogólniona zmienna położeniowa jest określona wzdłuż pętli, czyli jednowymiarowego tworu w przestrzeni. Uogólniony pęd jest strumieniem uogólnionego pola elektrycznego, przepływającym przez powierzchnię rozciągniętą przez pętlę. Nowe zmienne są zatem dużo bardziej „nielokalne" niż stare – w przeciwieństwie do metryki i krzywizny, które są określone punktowo, są to w pewnym sensie wielkości jedno- bądź dwuwymiarowe. Budzi to więc nadzieję, że uda się

uniknąć nieskończoności, które przecież są związane z nieskończenie małymi wielkościami.

W ramach grawitacji pętlowej można spodziewać się ciekawych wyników. Struktura przestrzeni w najmniejszych skalach wydaje się rzeczywiście dyskretna – analogicznie do dyskretnych wartości energii w widmach atomowych. Dyskretność ta ma oczywiście bezpośredni związek z pętlami. Funkcje falowe w ramach teorii pętli są zbudowane z tak zwanych sieci spinowych, czyli grafów, do których części przyporządkowuje się spiny całkowite lub połówkowe. Na rysunku 5.6 przedstawiono przykładową sieć spinową, składającą się z czterech części.



5.6. Przykład sieci spinowej; symbole od P_1 do P_4 oznaczają punkty przebicia części sieci przez powierzchnię S

Widoczna jest tam również powierzchnia *S*, którą przebija sieć spinowa. Zgodnie z formalizmem teorii pętli powierzchnie występują wyłącznie w formie wartości dyskretnych – są one dosłownie „skwantyzowane". Oczywiście jest to związane bezpośrednio ze skończoną liczbą punktów przecięcia sieci spinowej z powierzchnią i kwantyzacją spinu, który może przyjmować wyłącznie wartości całkowite lub połówkowe (zob. część 2). Dyskretna natura zmiennych pętlowych znajduje swoje odzwierciedlenie w dyskretnej naturze powierzchni. Istnieje elementarny „kwant

powierzchniowy” o wielkości odpowiadającej powierzchni Plancka (długość Plancka do kwadratu). Powierzchnie nie są jednak tylko wielokrotnościami kwantu powierzchniowego. Wraz ze zwiększającą się powierzchnią spektrum możliwych wartości zawęża się, a co za tym idzie – kwantyzacja powierzchni staje się istotna tylko w przypadku bardzo niewielkich powierzchni. Widać przez to wyraźnie, jak w konkretnym odniesieniu do grawitacji kwantowej potwierdzają się pierwotne hipotezy o istnieniu najmniejszej długości, należące do Bronsteina i innych.

Idea sieci spinowych pochodzi od Penrose’a, który wprowadził ją w innym kontekście na początku lat siedemdziesiątych XX wieku. Jego założenie było jednak podobne, a mianowicie stworzenie fundamentalnej teorii, która byłaby niezależna od tła. Pozwólmy dojść do głosu samemu autorowi:

> Mój pomysł polega na tym, aby stworzyć jednocześnie czasoprzestrzeń i mechanikę kwantową – w oparciu o zasady *kombinatoryczne* (...). Punktem wyjścia będą moment pędu, który posiada widmo *dyskretne*, oraz zasady wzajemnego powiązania momentów pędu, a celem – informacja, czy na tej podstawie można sformułować pojęcie *przestrzeni*.

Stworzenie samej przestrzeni z niewielkich dyskretnych jednostek bez uprzedniego założenia jej istnienia to bardzo ambitne przedsięwzięcie, które miałoby szanse powodzenia w ramach teorii pętli. W ujęciu tym nie brakuje jednak nierozwiązanych problemów, które w pewnym sensie uzupełniają te pochodzące z geometrodynamiki. Ujęcie to pozwala wprawdzie lepiej zrozumieć dyskretną strukturę mikrokosmosu, a jego strona matematyczna jest bardziej przejrzysta niż w przypadku geometrodynamiki, ale kwestia odnalezienia zaginionego parametru czasu, a wraz z nim fizyki klasycznej jako przypadku granicznego nadal pozostaje nierozstrzygnięta. Ponadto istnieje kilka problemów natury formalnej, które łączą się w tym ujęciu z odpowiednikiem równania Wheelera-DeWitta. Urok teorii pętli polega na ujawnieniu bardzo rozległych konsekwencji działania grawitacji kwantowej.

Czy dzięki teorii pętli można rozwiązać kilka problemów dotyczących pochodzenia wszechświata i czarnych dziur? W tym miejscu pragniemy skoncentrować się na czarnych dziurach, ponieważ w kolejnej części do głosu dojdzie ujęcie kosmologiczne.

Pod koniec części 3 dowiedzieliśmy się, że czarne dziury emitują promieniowanie, posiadają entropię i spełniają wymogi zasad termodynamiki. W związku z tym istnieją dwa problemy, które rozwiązać może tylko teoria grawitacji kwantowej. Pierwszy z nich dotyczy końcowej fazy parowania czarnych dziur, a drugi – mikroskopowej interpretacji entropii według Boltzmanna, czyli wyprowadzenia wzoru Bekensteina i Hawkinga, zgodnie z którym entropia jest proporcjonalna do powierzchni. Pierwszego problemu nie można na razie w żaden sposób wyjaśnić, natomiast w kwestii entropii ciekawe hipotezy przedstawia grawitacja pętlowa. Według Johna Wheelera, korzystając z heurystyki, można stworzyć obraz entropii: najpierw należy podzielić powierzchnię czarnej dziury (czyli powierzchnię horyzontu zdarzeń) na niewielkie pola o wielkości odpowiadającej powierzchni Plancka, a następnie pokryć je bitami, czyli wartościami 0 lub 1. Rysunek 5.7 jest przykładem takiego wyobrażenia.

5.7. Bity są rozmieszczone na powierzchni horyzontu czarnej dziury, a jeden bit odpowiada jednostce powierzchni Plancka

Entropia jest miarą ilości możliwości rozmieszczenia bitów na tej powierzchni. Ilość ta (a zarazem entropia) będzie rosła wraz z powierzchnią dziury. Z jakimi stopniami swobody mamy do czynienia w przypadku bitów? Mikroskopijnymi stopniami swobody gazu są na przykład położenia i prędkości molekuły gazu (zob. część 3). Entropia jest w tym przypadku miarą ilości możliwości wyboru tych położeń i prędkości do określonego makrostanu (określonego na przykład przez objętość i temperaturę). Teoria pętli proponuje, aby wyznaczyć mikroskopijne stopnie swobody czarnej dziury w grawitacji kwantowej na podstawie punktów przebicia, które sieć spinowa może tworzyć wraz z horyzontem zdarzeń. Po kilku nieudanych próbach ten skomplikowany problem kombinatoryczny udało się w końcu rozwiązać. Okazuje się bowiem, że entropia wyznaczona na podstawie wspomnianych sieci spinowych jest proporcjonalna do horyzontu zdarzeń. A zatem wspaniały triumf teorii pętli? Nie całkiem. Teoria pętli zawiera jeszcze jeden nowy parametr, a mianowicie liczbę, którą określa się mianem parametru Barbero-Immirziego. Jest sprawą otwartą, czy ze względu na podstawy

teoretyczne uda się znaleźć wartość doskonałą. Entropia jest proporcjonalna do tego nieznanego parametru. Oczywiście można również zabrać się do sprawy od końca i po prostu ogłosić, że wzór na entropię według Bekensteina i Hawkinga *musi* dać wynik. Następnie potwierdzi to parametr Barbero-Immirziego. Znaleziona w ten sposób wartość jest wprawdzie numeryczna, ale nie wyjaśnia niczego. Potrzebna jest więc druga możliwa do przetestowania prognoza teorii, aby przekonać się, czy grawitacja pętlowa sprawdza się z wyznaczonym w ten sposób parametrem. Niestety obecnie prognoza taka nie istnieje. Teoria pętli przedstawia jednak jasno, jak przynajmniej w teorii zinterpretować i policzyć mikrostany. W kolejnym rozdziale poznamy odrębny punkt widzenia.

Teoria strun

Ostatnie rozdziały były poświęcone kwantyzacji ogólnej teorii względności. Wyobrażenie, które za tym stoi, dotyczy sformułowania teorii grawitacji kwantowej bez konieczności troszczenia się o inne oddziaływania (silne i elektrosłabe), które pojawiają się dodatkowo w formie modelu standardowego fizyki cząstek elementarnych.

Kwestią sporną pozostaje nadal, czy taka bezpośrednia kwantyzacja grawitacji w ogóle jest możliwa. Alternatywne ujęcie zakłada zatem sformułowanie bezpośrednio uogólnionej teorii kwantowej *wszystkich* oddziaływań. Grawitacja kwantowa będzie w takim wypadku zaledwie jednym z aspektów tej teorii, który ujawni się, gdy w obrębie odpowiedniego przypadku granicznego poszczególne oddziaływania będzie można rozróżnić. Ujęcie takie – zresztą całkowicie unikatowe – nazywa się teorią strun.

Teoria strun może pochwalić się interesującą i burzliwą historią. W latach sześćdziesiątych XX wieku kwestia struktury składników jądra (proton, neutron) i odpowiednich silnych oddziaływań pozostawała nierozwiązana. Przeprowadzono wiele eksperymentów rozproszeniowych, zapoczątkowanych przez Rutherforda w 1911 roku, aby odkryć tę

strukturę. Można na przykład przy dużych energiach zderzać ze sobą protony i badać powstałe w wyniku tego cząstki. Wiele danych pochodzących z eksperymentów wyjaśniano, przyjmując, że w przypadku cząstek dochodziło do różnego wzbudzenia drgającej struny. Z uwagi na to, że słowo „struna" to po angielsku „string", mówiono o modelu strunowym oddziaływań silnych. W tej książce również będziemy posługiwać się tym utartym pojęciem.

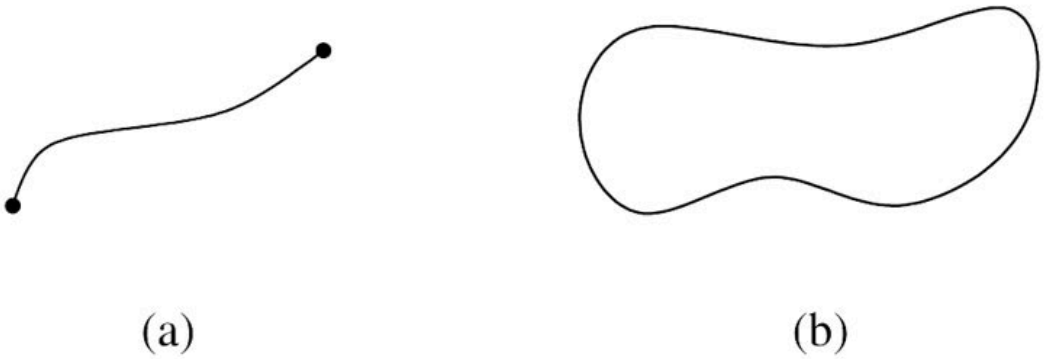
Czy w jądrach atomów kryją się zatem maleńkie struny gitarowe? Dalsze eksperymenty nie mogły podtrzymać tego obrazu. Okazało się bowiem, że cząstki jądra składają się z punktowych obiektów zwanych kwarkami, których oddziaływanie z powodzeniem opisuje chromodynamika kwantowa, czyli teoria cechowania oddziaływania silnego (zob. część 2). Wydawało się więc, że model strunowy już się nie rozwinie i dokona swojego żywota jako jedno z wielu niepowodzeń w historii nauki.

Sytuacja zmieniła się w 1974 roku, gdy zespół fizyków pokazał, że w widmie struny znajduje się bezmasowa cząstka o spinie = 2. Cząstkę tę poznaliśmy już wcześniej – to grawiton, który odpowiada za oddziaływanie grawitacyjne. Po przyjęciu ogólnych założeń można pokazać, że grawiton jest jedyną bezmasową cząstką o spinie = 2. Stwierdzenie to wywołało niemałe poruszenie wśród fizyków. Czy model strunowy jest w stanie dostarczyć nam teorię grawitacji kwantowej?

Model strunowy zawiera jeden wolny parametr o wymiarze długości – a dokładniej: długości struny. W modelu strunowym oddziaływań silnych długość ta wynika w sposób naturalny z wielkości jądra, a zatem powinna wynosić około 10^{-13} centymetra. Jeśli uznamy model za teorię kwantową, która uwzględnia również grawitację, to długość struny powinna odpowiadać długości Plancka, a więc około 10^{-33} centymetra, czyli 20 jednostek mniej. Oczywiście w przypadku tak niewielkich strun znika problem sprzeczności z eksperymentem. W skalach długości, odczytywanych przez obecne akceleratory, struny te przypominają raczej punkty i dlatego trudno je rozpoznać.

W przeciwieństwie do strun gitarowych nie są one na stałe przymocowane, a zatem nie mają stałej długości zdefiniowanej przez warunki brzegowe. Charakteryzują się jednak ogromnym naprężeniem, które jest niezbędne do tego, aby pozostały one mikroskopijne, czyli posiadały długość Plancka.

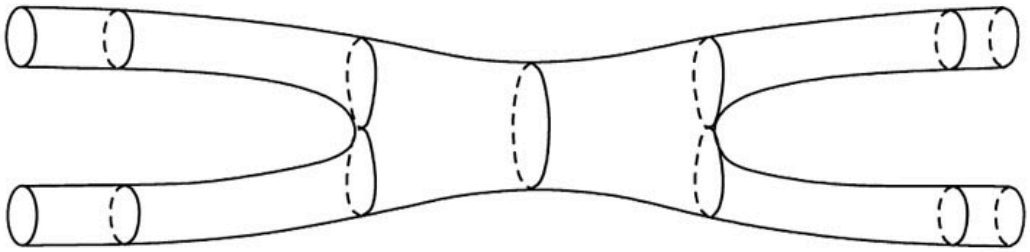
Istnieją dwa rodzaje strun. W przypadku, gdy ich końce są swobodne, mówimy o strunach otwartych, a gdy złączone – o strunach zamkniętych. Rysunek 5.8 przedstawia oba te warianty.



5.8. Struna otwarta (a) i zamknięta (b)

W teorii strun nie występują obiekty nieustrukturyzowane („punktowe”), a zatem wszystkie cząstki, niezależnie od tego, czy są to elektrony, czy protony, są wyłącznie efektem wzbudzenia strun fundamentalnych. Podobnie jak w przypadku grawitacji pętlowej budzi to nadzieję na to, że dywergencje kwantowej teorii pola znikną, ponieważ struny nie oddziałują tylko w pojedynczych punktach czasoprzestrzeni, lecz są rozproszone w wielu punktach. Nadzieje te w dużej mierze się spełniły. Na rysunku 5.9 przedstawiono przykład oddziaływania dwóch strun zamkniętych. Czas biegnie od lewej do prawej strony; napływają dwie struny zamknięte, oddziałują między sobą i odpływają w przeciwnych kierunkach. W czasoprzestrzeni oddziaływanie opisane jest więc za pomocą „podwójnych spodni”, których nie wyróżnia

pojedynczy punkt. W związku z tym nieskończoności połączone z takimi punktami w ogóle nie występują.



5.9. Oddziaływanie dwóch strun zamkniętych

Można zatem przedstawić ogólne stwierdzenie, że teoria strun bez grawitacji nie jest spójna. Edward Witten, jeden z głównych teoretyków w zakresie teorii strun, poszedł nawet o krok dalej, twierdząc, że istnienie grawitacji jest jedną z najlepszych prognoz teorii strun. Jeśli więc w wyniku przerażenia książka ta wypadnie Państwu z rąk i spadnie na podłogę, to będzie to efekt teorii strun.

Zgodnie z tym, co napisano wcześniej, teoria strun wprowadza do fizyki nowy parametr – długość struny. Ponadto są jeszcze starzy znajomi: *h* (kwant działania) oraz *c* (prędkość światła). W zasadzie na podstawie teorii strun powinna zatem istnieć możliwość obliczenia wszystkich innych stałych w fizyce (masy, stałe sprzężenia, a także stała grawitacyjna). Z całą pewnością byłaby to duża zaleta tej teorii, gdyż wspomniana wyżej kwantyzacja teorii względności zdaje się nie zapewniać takich wyliczeń z uwagi na nieuwzględnioną unifikację. Z góry należy jednak podkreślić zwrot „w zasadzie”, ponieważ nasza wiedza w zakresie teorii strun jest jeszcze niewystarczająca, aby przeprowadzić takie wyliczenia.

Im większe naprężenie struny, tym więcej energii jest potrzebne, aby wprawić ją w ruch. Problem w tym, że naprężenie strun fundamentalnych jest ogromne. Przez to, że ich długość odpowiada długości Plancka, energie wzbudzenia strun powinny odpowiadać energii Plancka (według

Einsteina: masa Plancka razy c^2). Należałoby zatem oczekiwać, że cząstki, które pojawiają się w wyniku wzbudzenia strun, mają masy odpowiadające przynajmniej masie Plancka. Jak to możliwe? Wszystkie znalezione cząstki mają masy znacznie mniejsze od masy Plancka. Czy w związku z tym teoria strun jest jednak nic niewarta?

W tym miejscu decydującą rolę odgrywają aspekty kwantowe. Podobnie jak w przypadku efektu Casimira (zob. część 2) teoria kwantowa również pozwala na ujemne energie. W odniesieniu do podstawowego drgania struny okazuje się, że taka ujemna energia może zniwelować oczekiwaną wartość masy o wymiarach Plancka. W razie całkowitego zniesienia tej wartości powstaną cząstki o zerowej masie spoczynkowej (na przykład foton i grawiton). Jeżeli masa spoczynkowa otrzyma niewielką wartość dodatnią, powstaną pozostałe cząstki, takie jak na przykład elektron. Obserwowany świat otrzymaliśmy zatem w wyniku podstawowego drgania struny. Skłoniło to teoretyka w zakresie teorii strun i autora bestsellerów Briana Greene'a do stworzenia wizji wszechświata jako kosmicznej symfonii. (Najwyraźniej nie pomyślał on o możliwych dysonansach).

Wyższe drgania strun opisują z kolei cząstki o masie Plancka i cięższe. Możliwość ich bezpośredniej obserwacji dawałyby wyłącznie akceleratory o rozmiarach galaktyki, co jest rzecz jasna nierealnym przedsięwzięciem. Drgania wyższe można zobaczyć co najwyżej pośrednio, ponieważ są zawarte automatycznie w strukturze teorii strun.

Zbudowanie spójnej teorii strun w ramach teorii kwantowej nie jest łatwym zadaniem. W pierwotnej teorii strun okazało się, że możliwe jest to tylko wtedy, gdy struna porusza się w czasoprzestrzeni o 26 wymiarach! To trochę więcej niż zaobserwowane cztery wymiary czasoprzestrzeni. O dodatkowe problemy postarała się niezwykła cząstka występująca w widmie, zwana tachionem.

Tachiony są cząstkami hipotetycznymi, poruszającymi się z prędkością nadświetlną. Mimo że może wydawać się inaczej, to jednak ich egzystencja jest zgodna ze szczególną teorią względności, gdyż mogą one poruszać się *wyłącznie* z prędkością nadświetlną, a więc żadne przejście

między prędkościami mniejszymi i większymi od *c* nie jest możliwe. Tachiony sprawiają problemy, gdy uwzględnia się teorię kwantową. Próżnia może stać się wtedy niestabilna i rozpaść się na tachiony, czego rzecz jasna nie zaobserwowano.

O rozwiązanie problemu tachionów troszczy się dodatkowa symetria, z którą spotkaliśmy się już w części 2, czyli tak zwana supersymetria (SUSY). Jest to symetria, która może istnieć pod warunkiem uwzględnienia spinu cząstek. SUSY zakłada, że w przyrodzie cząstki zawsze łączą się w pary, przy czym cząstki pary („superpartnerzy") różnią się od siebie spinem połówkowym. Zgodnie z powyższym superpartnerem elektronu (o spinie = ½) jest na przykład selektron (o spinie = 0). W przypadku fotonu (spin = 1) jest to fotino (spin = ½), a grawitonu (spin = 2) – grawitino (spin = 3/2). Jeden z partnerów jest zatem zawsze fermionem, a drugi – bozonem. Z uwagi na to, że bozony przekazują oddziaływania (foton – elektromagnetyczne, a grawiton – grawitacyjne), SUSY łączy w pewnym sensie materię i oddziaływania.

Źródłem wzbudzeń w pierwotnej teorii strun były wyłącznie bozony, stąd też mówi się o tak zwanej strunie bozonowej. Z oczywistych powodów chciałoby się jednak, aby teoria strun uwzględniała również fermiony. Na początku lat siedemdziesiątych XX wieku podczas próby doimplementowania fermionów automatycznie natknięto się na SUSY, a więc bez SUSY nie można zbudować struny z fermionami. W przypadku uwzględnienia fermionów, a tym samym także SUSY, mówi się o teorii superstrun. (Dla uproszczenia w dalszej części tekstu będziemy mówić głównie o teorii strun). Dopiero w późniejszym okresie zauważono, że SUSY można również przenieść na grunt zwykłej fizyki cząstek elementarnych, gdzie będzie dostarczać uogólnione (dotąd niesprawdzone) modele.

Teoria superstrun unika tachionu. Czy unika ona też nazbyt licznych wymiarów czasoprzestrzeni struny bozonowej? Tylko częściowo. Teorię tę można sformułować w sposób spójny wyłącznie w dziesięciu wymiarach czasoprzestrzeni – zawsze to lepiej niż w 26, ale mimo wszystko to wciąż

więcej niż oczekiwane cztery wymiary. Jak rozumieć te dodatkowe wymiary? W części 1 omawialiśmy starą koncepcję Kaluzy i Kleina, według której można rozszerzyć ogólną teorię względności o piąty wymiar poprzez połączenie teorii względności *oraz* elektromagnetyzmu w czterech wymiarach. Piąty wymiar musiałby być na tyle mały, aby uniknąć sprzeczności z eksperymentem.

Pomysł Kaluzy i Kleina przeżywa w ramach teorii strun prawdziwe odrodzenie. Oferuje on bowiem klucz do zrozumienia nieuniknionych wyższych wymiarów. Należy sobie zatem wyobrazić, że sześć z dziesięciu wymiarów czasoprzestrzeni jest zwiniętych w niewielkie twory, które w każdym punkcie przypięte są do czterech makroskopijnych wymiarów. Z uwagi na to, że rozmiar tych dodatkowych wymiarów jest zbliżony do długości Plancka, staje się jasne, dlaczego nie zdawaliśmy sobie sprawy z ich istnienia: po prostu są one zbyt małe, abyśmy mogli je zauważyć. (W następnym rozdziale poznamy jeszcze ogólniejszy sposób na wyjaśnienie wyższych wymiarów).

Struktura geometryczna wyższych wymiarów jest skomplikowana. W matematyce twory te nazywa się przestrzeniami Calabiego-Yau. Posiadają one wiele dziur, które mają wpływ na wzorzec drgań struny. W ten sposób, w przypadku przestrzeni z trzema dziurami, można by było założyć bezpośredni związek z trzema rodzinami fizyki cząstek elementarnych (zob. część 2) – ekscytująca wizja. Jak dotąd niestety nie jest znana żadna zasada, która opierałaby się na przestrzeniach Calabiego-Yau. To ogólny problem, ponieważ istnieje stanowczo zbyt wiele wariantów budowy tych przestrzeni (nie wspominając nawet o możliwościach, które omówimy w następnym rozdziale), aby móc formułować jednoznaczne prognozy dla fizyki czterowymiarowej. Obliczono nawet, że istnieje około 10^{500} możliwych wariantów, aby przejść od fundamentalnych dziesięciu do zaobserwowanych czterech wymiarów. To, co to oznacza w kontekście trafności prognoz danej teorii, trzeba jeszcze wyjaśnić.

W rzeczywistości supersymetrię można przenieść na grunt teorii strun na pięć różnych sposobów. Dwa z nich bardzo zręcznie łączą strunę supersymetryczną z bozonową – z tego powodu nazywa się ją „struną heterotyczną". Struna ta była główną bohaterką wydarzenia, które nieco górnolotnie nazywa się „pierwszą rewolucją w teorii strun". Opublikowane w 1984 roku prace teoretyków z zakresu teorii strun – Michaela Greena i Johna Schwarza – pokazały, że struny heterotyczne oferują względnie realistyczne modele fizyki cząstek elementarnych, a w szczególności potrafią opisać istotne w tym obszarze łamanie parzystości (czyli fakt, że procesy przebiegają inaczej, jeśli obserwuje się je w lustrze). Teoria strun mogłaby zatem mieć rzeczywiście coś wspólnego z obserwowanym światem. Po długim okresie stagnacji teoria zyskała żywe zainteresowanie w społeczności fizyków cząstek elementarnych.

Do „drugiej rewolucji w teorii strun" doszło 11 lat później. W 1995 roku Edward Witten ogłosił na konferencji naukowej, zorganizowanej na temat teorii strun, że pięć znanych teorii nie jest od siebie niezależnych, lecz w rzeczywistości wiąże się ze sobą siatką tak zwanych dualności. Zgodnie z powyższym przypadek graniczny słabego sprzężenia strun jednej teorii jest równy przypadkowi granicznemu silnego sprzężenia drugiej. Ponadto udało się stwierdzić, że w rozwiązaniach teorii strun występują nie tylko struny, czyli nie tylko twory jednowymiarowe. W rzeczywistości istnieje wiele innych obiektów o różnych wymiarach (dwóch, trzech lub więcej), które określa się zbiorczo jako „brany" (nazwa pochodzi od drugiej sylaby słowa „membrana"). Ostatecznie jednak nadal najważniejsze są struny, ponieważ masy bran w przypadku niewielkich sprzężeń (tak zwana teoria perturbacji) są duże, a przez to nieistotne.

Druga rewolucja w teorii strun wiązała się również z innym odkryciem – w formalizmie pojawił się tajemniczy jedenasty wymiar. Ma to związek z tym, że udało się wyjść poza granice teorii perturbacji. Niektóre struny są zatem w rzeczywistości dwuwymiarowymi membranami, przy czym jeden z tych wymiarów wskazuje na jedenasty wymiar, którego w teorii perturbacji (czyli przy niewielkim sprzężeniu) nie można zidentyfikować.

Przed 1995 rokiem we wspomnianych dziesięciu wymiarach membrana ta rzeczywiście wygląda jak struna.

Zdaje się więc, że istnieje zbiorcza teoria określona przez jedenaście wymiarów czasoprzestrzeni, w ramach której pięć pierwotnych teorii strun stanowi różne przypadki graniczne. Mimo że superteoria ta nie jest jeszcze dobrze zbadana, otrzymała już swoją nazwę. Określa się ją mianem *m-teorii*. Co oznacza *m*? W tej kwestii także istnieją różne wyobrażenia. Być może pochodzi ono od angielskiego słowa „mystery" (zagadka, tajemnica), albo od słowa „matka", a może od czegoś innego. Tajemnica dotycząca tej teorii nie została jeszcze wyjaśniona. Być może uczyni to kolejna rewolucja w teorii strun.

Odkrycie Wittena z 1995 roku dotyczyło przede wszystkim związków między różnymi teoriami strun (dualności). Inny rodzaj dualności jest związany bezpośrednio z dodatkowymi wymiarami i skończoną długością struny. Wyobraźmy sobie dodatkowy wymiar przestrzeni, który jest zwinięty w okrąg o promieniu *R*. Z uwagi na skończoną długość struny może ona rozwijać się na różne sposoby. Może poruszać się jednostajnie, może drgać, bądź też może owinąć się wokół dodatkowego wymiaru. Okazuje się bowiem, że ruch jednostajny w przypadku dużego okręgu odpowiada owijaniu się wokół małego okręgu i odwrotnie. Z perspektywy samego operacjonizmu nie można zatem odróżnić wszechświata, w którym dodatkowy wymiar w kształcie okręgu ma promień *R*, od odpowiedniego wszechświata o promieniu proporcjonalnym do 1/*R*. Sposób postrzegania zależy od tego, czy do pomiaru długości wykorzystano strunę poruszającą się jednostajnie czy zwiniętą. Wniosek taki nie może oczywiście wynikać z teorii cząstek punktowych. Nie zaskakuje również fakt, że zasada nieoznaczoności w mechanice kwantowej została uogólniona. Rozciągnięta struna nie pozwala po prostu na odczytanie struktury mniejszej niż długość struny. W przypadku zwiększenia nieoznaczoności pędu nieoznaczoność położenia najpierw się zmniejszy, potem jednak, po zbliżeniu się do długości struny – ponownie wzrośnie.

W poprzednich rozdziałach stale podkreślaliśmy, jak istotne są poszukiwania koncepcji grawitacji kwantowej, która nie uwzględniałaby stałego tła. Teoria strun przedstawia nam z kolei obraz niewielkich strun i innych tworów poruszających się w dziesięcio- lub jedenastowymiarowej czasoprzestrzeni, która sama nie została skwantyzowana. W rzeczywistości widać zatem, że teoria strun nie jest jeszcze kompletna. W fundamentalnej koncepcji m-teorii nie będzie już miejsca dla czasoprzestrzeni. Kwestia tego, czy tak jak w teorii kanonicznej zostanie zachowana przestrzeń, czy również i jej zabraknie, pozostaje nadal otwarta. Samą czasoprzestrzeń powinno dać się z grubsza zbudować ze strun i bran. Dalszy rozwój sytuacji powinien przybrać ciekawy obrót. Całe środowisko fizyków ubolewa nad faktem, że jak dotąd nie sformułowano żadnej fundamentalnej zasady, takiej jak na przykład zasada równoważności, na podstawie której powstałaby teoria strun (bądź m-teoria). Czy w ogóle istnieje taka zasada?

W poprzednim rozdziale przekonaliśmy się, że pętlowa grawitacja kwantowa w pewnym sensie pozwala na makroskopowe wyjaśnienie entropii czarnych dziur. Było to możliwe dzięki policzeniu stopni swobody powiązanych z sieciami spinowymi. Jednak z uwagi na obecny w formalizmie parametr Barbero-Immirziego rozumowanie to nie było pozbawione luk. Jak wygląda sytuacja w przypadku teorii strun? W jej kontekście również mówi się o mikroskopijnych stopniach swobody w formie strun i bran. Bekensteinowi i Hawkingowi udało się pokazać, że odnalezienie wzoru pozwalającego na prawidłowe określenie entropii jest rzeczywiście możliwe dzięki odpowiednio zdefiniowanemu zliczeniu stanów branowych. Niestety istnieje jeden haczyk. Entropia jest wprawdzie poprawna, jednak przedstawione czarne dziury nie są zgodne z zaobserwowanymi czarnymi dziurami, jak na przykład z supermasywną dziurą znajdującą się w centrum naszej Drogi Mlecznej. Są to raczej dziury o dodatkowych ładunkach strunowych (uogólnionych ładunkach elektrycznych), w przypadku których niezbędne jest dodatkowo wystąpienie odpowiedniej zależności między nimi a masą dziury. Mimo to pierwszy krok został już zrobiony.

Zgodnie z tym, co pisaliśmy już wcześniej, wciąż brakuje możliwości eksperymentalnej weryfikacji teorii strun. Ze względu na ogromne energie eksperymenty przeprowadzane bezpośrednio w akceleratorach nie mają szans powodzenia. Jedynym wyjściem byłoby powstanie opisanej w następnym rozdziale kryjówki. W zasadzie powinna istnieć możliwość obliczenia mas wszystkich znanych nam cząstek oraz stałych sprzężenia modelu standardowego fizyki cząstek elementarnych, lecz aktualnie nie da się tego zrealizować. Winę za to ponosi niezrozumiale wysoka liczba możliwości przejścia od dziesięciu bądź jedenastu wymiarów czasoprzestrzeni, pochodzących z teorii fundamentalnej, do świata czterowymiarowego, wynikającego z naszego doświadczenia. Jednak pomijając to, teoria, również w przypadku energii dużo niższych od energii Plancka, powinna prognozować niewielkie odchylenia od teorii względności, jak na przykład naruszenie zasady równoważności bądź szczególnej teorii względności, obowiązującej lokalnie w swobodnie spadającym układzie. Eksperymenty, które mają to potwierdzić, trwają i nie ograniczają się wyłącznie do wspomnianych w części 1 misji kosmicznych MICROSCOPE i STEP. Poszukiwania te przypominają jednak błądzenie w ciemności, ponieważ dokładne prognozy co do stopnia takich odchyleń nie istnieją.

Zarówno teoria strun, jak i pętli opierają się na wyobrażeniu o istnieniu najmniejszej długości, nie rozstrzygając przy tym tego, czy istnieje ona tylko w rozumieniu operacjonizmu, czy też nie. Tym samym pożegnano się z pomysłem ciągłego opisywania przestrzeni. Czy dzięki eksperymentom można dowieść istnienia ziarnistej struktury przestrzeni? Nie jest to wcale pozbawione sensu, choć na pierwszy rzut oka mogłoby się tak wydawać. Sposób rozprzestrzeniania się promieniowania elektromagnetycznego bądź wysokoenergetycznych cząstek kosmicznych odróżnia dyskretną strukturę przestrzeni od ciągłej. W przypadku, gdy w wyniku zdarzenia, do którego doszło w kosmosie, emitowane jest promieniowanie elektromagnetyczne, w przestrzeni o strukturze ziarnistej różne długości fali mogą charakteryzować się prędkościami odbiegającymi od siebie w niewielkim stopniu. Jedno zdarzenie w kosmosie może mieć

zatem różny czas dotarcia do Ziemi. W takich przypadkach mówi się najczęściej o tak zwanych rozbłyskach promieni gamma, które mogły powstać w wyniku połączenia obiektów kompaktowych (czarne dziury, gwiazdy neutronowe). Mimo prowadzenia intensywnych poszukiwań nie natrafiono do tej pory na żaden ślad – czasy dotarcia do Ziemi różnych długości fali są z uwzględnieniem niedokładności pomiarowych takie same. Poszukiwania trwają.

Wszechświat na branie

W poprzednim rozdziale przekonaliśmy się, że teoria strun wymaga większej liczby wymiarów przestrzeni – zgodnie z powyższym czasoprzestrzeń posiada zatem dziesięć lub jedenaście wymiarów. Z uwagi na to, że do tej pory nie dostrzegliśmy tych dodatkowych wymiarów, powinny one być niezwykle małe i wynosić mniej niż 10^{-17} centymetra, co odpowiada najmniejszej odległości zarejestrowanej przez dostępne nam obecnie akceleratory. Oznacza to ożywienie starych idei Kaluzy i Kleina, które poznaliśmy pod koniec pierwszej części tej książki. Tamta teoria dotyczyła integracji grawitacji z elektrodynamiką poprzez koncepcję podstawowej teorii grawitacji w pięciu wymiarach. W tym wypadku chodzi natomiast o integrację wszystkich oddziaływań poprzez teorię strun (bądź m-teorię) w dziesięciu lub jedenastu wymiarach.

Pod koniec lat dziewięćdziesiątych XX wieku okazało się bowiem, że rzeczywiście istnieje o wiele więcej możliwości integracji dodatkowych wymiarów z teorią, niż wynika to ze starej koncepcji Kaluzy i Kleina i jej bezpośrednich rozszerzeń. Zgodnie z tak zwanym modelem ADD (powstałym w 1998 roku, którego nazwa pochodzi od pierwszych liter nazwisk trzech fizyków: Arkaniego-Hameda, Dimopoulosa oraz Dvaliego) dodatkowe wymiary nie muszą być wcale tak mikroskopijne – autorzy spekulują, że ich wielkość może sięgać jednego milimetra. Jak to możliwe? Czyż nie stwierdziliśmy przed chwilą, że akceleratory nie zarejestrowały obecności żadnych dodatkowych wymiarów aż do 10^{-17} centymetra? Tak, to prawda, o ile uwzględniamy wszystkie

oddziaływania. W teorii strun istnieją jednak również modele opierające się na branach (zob. wyżej), w których wszystkie oddziaływania *poza* grawitacją ograniczają się właśnie do tych bran i nie mogą mieć wpływu na pozostałe wymiary. Grawitacja jako wyraz geometrii czasu i przestrzeni jest oczywiście obecna wszędzie, ponieważ nie może być niezależna od czasoprzestrzeni.

Wyobraźmy sobie, że nasz czterowymiarowy wszechświat (trzy wymiary przestrzeni i jeden wymiar czasu) odpowiada takiej branie. Ponadto dla uproszczenia przyjmijmy również, że istnieje tylko jeden dodatkowy, a zatem piąty wymiar (cały świat miałby więc cztery wymiary przestrzeni i jeden wymiar czasu). Jeśli tylko grawitacja przenika przez wszystkie pięć wymiarów, to piąty wymiar można namierzyć wyłącznie przez pomiar grawitacji. Dodatkowy wymiar zmienia na przykład prawo grawitacji Newtona, którego nie można sprawdzić dla mikroskopijnych odległości, ponieważ trudno jest dokonać wiarygodnego pomiaru między dwiema tak blisko siebie położonymi masami. Pod koniec lat dziewięćdziesiątych sprawdzono prawo grawitacji tylko dla odległości około jednego milimetra, a zatem dodatkowe wymiary mogłyby mieć maksymalnie taką właśnie wielkość. (W późniejszych latach udało się potwierdzić działanie prawa grawitacji dla odległości około 100 mikrometrów, co jednak niewiele zmienia).

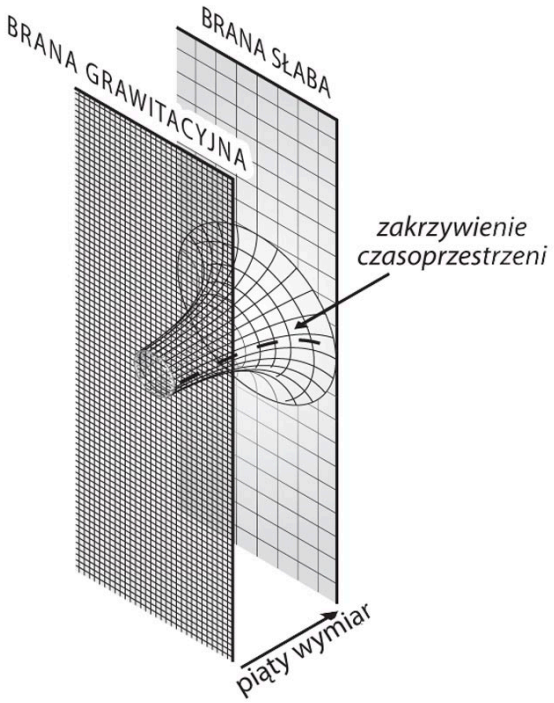
Głównym powodem opracowania modelu ADD była chęć rozwiązania tak zwanego problemu hierarchii w fizyce cząstek elementarnych. Pod owym pojęciem kryje się fakt związany z tym, że skala energii oddziaływania słabego jest wielokrotnie mniejsza niż energia Plancka, charakterystyczna dla grawitacji kwantowej. Różnica ta wynosi około 10^{16}, a więc 16 potęg liczby 10, i jako taka nie jest problemem samym w sobie. We wszechświecie znanych jest nam przecież wiele struktur o diametralnie różnych skalach długości, na przykład człowiek, który jest o 26 potęg liczby 10 mniejszy od wszechświata i nikogo to nie dziwi. Problem hierarchii staje się prawdziwym problemem dopiero ze względu na słynną cząstkę Higgsa, która jest odpowiedzialna za oddziaływanie

słabe i w szczególności za wielkość masy cząstek. Zgodnie z formalizmem modelu standardowego należałoby bowiem oczekiwać, że masa cząstki Higgsa jest niestabilna i dąży do osiągnięcia wielkości odpowiadającej masie Plancka – co oczywiście nie jest prawdą.

Z uwagi na znacznie większe rozmiary wymiarów dodatkowych modelu ADD długość Plancka w teorii pięciowymiarowej mogłaby osiągnąć nawet jeden milimetr (zgodnie ze wspomnianym wyżej ograniczeniem wynikającym z weryfikacji prawa grawitacji) przy odpowiednio mniejszej masie Plancka. Oznaczałoby to wtedy, że masa Higgsa odpowiada mniej więcej masie Plancka, a problem hierarchii nie istnieje. Realny problem stałby się problemem pozornym, wynikającym z naszej ignorancji w kwestii istnienia piątego wymiaru. Oczywiście model ADD należy na razie uznawać jedynie za możliwą hipotezę, a nie rozwiązanie problemu, ponieważ samo istnienie dodatkowego wymiaru o pożądanych właściwościach wymaga wyjaśnienia, o które postarać mogłaby się wyłącznie teoria fundamentalna, jak na przykład teoria strun (do czego jeszcze nie jest gotowa). Zręczna hipoteza jest jednak często pierwszym krokiem do znalezienia rozwiązania.

Ponadto istnieją jeszcze inne ciekawe możliwości. W 1999 roku amerykańscy fizycy Lisa Randall oraz Raman Sundrum uraczyli nas swoim modelem, w ramach którego problem hierarchii można rozwiązać również w przypadku mikroskopijnych wymiarów dodatkowych. Model ten składa się z dwóch bran (każda z nich posiada po cztery wymiary czasoprzestrzeni) rozdzielonych piątym wymiarem. W odróżnieniu od modelu ADD koncepcja ta uwzględnia także pole grawitacyjne wytworzone przez obie brany. Z racji obecności dwóch bran konieczne było też rozwiązanie pięciowymiarowych równań Einsteina. Wynikiem tego jest tak zwana zakrzywiona (z angielskiego: *warped*) geometria. Inaczej niż w przypadku idei Kaluzy i Kleina oraz ich następców, metryka czterowymiarowej części czasoprzestrzeni (a tym samym długości, czasów i mas) zależy w tym przypadku od wymiarów dodatkowych. Wprawdzie czterowymiarowy świat, znajdujący się w stałej pozycji wzdłuż piątego

wymiaru, jest zawsze płaski (dotyczy to również obu bran), ale od tej pozycji zależą wszystkie wielkości geometryczne pozostałych czterech wymiarów. Z modelu Randall i Sundruma wynika silny („wykładniczy") spadek siły grawitacji w miarę przemieszczania się przez piąty wymiar od jednej brany do drugiej. Znajduje to swoje odzwierciedlenie na przykład w zależności między dokładnością pracy zegarka a położeniem w piątym wymiarze. Zależność tę przedstawia rysunek 5.10. W związku z tym pierwszą branę nazywa się braną grawitacyjną, a drugą – braną słabą. Według wyobrażenia Randall i Sundruma brana słaba jest obserwowanym przez nas czterowymiarowym światem, w którym obowiązuje model standardowy fizyki cząstek elementarnych wraz ze znanymi nam masami cząstek. Ich niewielki rozmiar w stosunku do masy Plancka wynika w tym modelu z tego, że siła grawitacji maleje wykładniczo od brany grawitacyjnej (gdzie wszystkie masy są takie same) do brany słabej. Ze względu na ten szybki spadek siły grawitacji odległość między branami wcale nie musi być taka duża – kilka długości Plancka wystarczy.



5.10. Spadek siły grawitacji w miarę przemieszczania się od brany grawitacyjnej do brany słabej

Model ten jest rzecz jasna również jedynie hipotezą i polega wyłącznie na przeformułowaniu problemu hierarchii. Ma on jednak swój urok, ponieważ w przypadku istnienia dwóch bran jedynym warunkiem koniecznym jest tylko ogólna teoria względności Einsteina w pięciu wymiarach. Ponadto model ten (podobnie jak model ADD i inne) można sprawdzić eksperymentalnie! Wspomniany już w części 2 akcelerator LHC, który zostanie uruchomiony w pobliżu Genewy, powinien wykryć ewentualnie istniejące wyższe wymiary. Ich istnienie w naszym świecie manifestują bowiem cząstki o dobrze określonych właściwościach – tak zwane cząstki Kaluzy-Kleina. Cząstka taka może powstać na przykład w wyniku zderzenia dwóch protonów, aby po chwili rozpaść się, tworząc parę złożoną z elektronu i pozytronu. W warunkach doświadczalnych proces taki będzie posiadał właściwości, które umożliwią jego jednoznaczną identyfikację. Modele te pozwolą nawet na utworzenie

w LHC niewielkich czarnych dziur, które będą tak mikroskopijne, że będzie można mówić o prawdziwych obiektach kwantowych. W wyniku emisji promieniowania Hawkinga po powstaniu czarnych dziur dojdzie do ich bezzwłocznego wyparowania (dzięki temu są one również nieszkodliwe). Zaobserwowanie tych niewielkich dziur także powinno być możliwe. Czy jednak przyroda wyświadczy nam tę przysługę i uraczy nas dodatkowymi wymiarami? Tego nie wiemy, ale przyszłość na pewno zapowiada się ciekawie.

Czy w czarnych dziurach gubią się informacje?

Ścisk przed aulą był ogromny. 21 kwietnia 2004 roku przed centrum konferencyjnym Towarzystwa Królewskiego w Dublinie (Royal Dublin Society) zebrał się tłum, aby posłuchać wykładu Stephena Hawkinga. Temat wykładu brzmiał następująco: „Paradoks informacyjny czarnych dziur". Hawking ogłosił, że zmienił swoją opinię, której hołdował od 1976 roku. Co w tym interesującego?

W części 2 tej książki dowiedzieliśmy się, że czarne dziury nie są tak naprawdę czarne, jeśli w ich opisie uwzględni się teorię kwantową. Czarne dziury emitują promieniowanie o temperaturze proporcjonalnej do kwantu działania h – jest to tak zwana temperatura Hawkinga. Z uwagi na to, że promieniowanie przenosi energię, jest ona pobierana z dziury, co powoduje spadek jej masy. Temperatura Hawkinga jest jednak odwrotnie proporcjonalna do masy dziury, co w miarę trwania emisji promieniowania prowadzi do dalszego jej nagrzewania. Po pewnym czasie dziura wyparowuje. Ale co po niej pozostaje?

Zgodnie z obliczeniami Hawkinga należałoby przede wszystkim oczekiwać, że pozostanie po niej tylko promieniowanie cieplne. Wiązałby się z tym jednak jeden problem. Czarna dziura może powstać z wielu różnych stanów początkowych. Zazwyczaj kolaps grawitacyjny masywnych gwiazd, których masa jest około trzykrotnie większa od masy Słońca, prowadzi do powstawania czarnych dziur. Dzieje się tak niezależnie od pozostałych właściwości takiej gwiazdy. Ponadto można sobie wyobrazić,

co stałoby się z obiektami o różnej zawartości informacyjnej, na przykład telewizorami bądź samochodami, po wrzuceniu ich do takiej dziury. Jeśli horyzont zdarzeń oddziela wnętrze czarnej dziury od strefy zewnętrznej, można założyć, że informacja trafi do wnętrza i pozostanie niewidoczna. Gdyby zatem po wyparowaniu, czyli po zaniku horyzontu, pozostało samo promieniowanie cieplne, oznaczałoby to utratę ogromnej ilości informacji: niezależnie od stanu początkowego stan końcowy byłby zawsze taki sam. Stałoby to w sprzeczności z determinizmem zarówno fizyki klasycznej, jak i fizyki kwantowej (w rozumieniu funkcji falowych, ewoluujących w sposób deterministyczny). Wykład Hawkinga w Dublinie dotyczył właśnie tego „paradoksu informacyjnego".

Prawdopodobnie w pierwszej chwili w ogóle nie dostrzegamy powagi całej sytuacji. Przecież w normalnym świecie poza czarnymi dziurami potrafimy zniszczyć różne informacje, na przykład paląc zapisaną kartkę papieru, czyż nie? Z praktycznego punktu widzenia: tak, informacja zostaje zniszczona. Zasadniczo jednak nie znika ona z naszego świata, mimo że jest dla nas niedostępna. Każde inne twierdzenie byłoby niezgodne z determinizmem. Przejście od informacji istotnej (zapisanej kartki papieru) do nieistotnej (spalonej kartki papieru wraz ze zmienionym stanem otoczenia) zostało opisane przez wzrost entropii, o którym mówi druga zasada termodynamiki (zob. część 3). W przypadku czarnej dziury rzeczywiście doszłoby do całkowitego zniszczenia informacji, gdyby cały układ (trzeba wyobrazić sobie odizolowaną czarną dziurę pozbawioną wszelkiego otoczenia) zawsze przeobrażał się wyłącznie w promieniowanie cieplne.

Pierwotnie Hawking był zdania, że informacja rzeczywiście zostaje całkowicie zniszczona. Z uwagi na to, że końcową fazę parowania musiałaby opisać nieistniejąca jeszcze grawitacja kwantowa, teoria taka miałaby zupełnie nowe, nieznane nam właściwości. Dość szybko pojawiły się głosy sprzeciwu wobec tej opinii. Wzór Hawkinga na obliczenie temperatury opiera się na kilku założeniach i przybliżeniach. Krytycy tego rozwiązania twierdzą, że gdyby wzór ograniczył się całkowicie do

grawitacji kwantowej, to promieniowanie nie byłoby wyłącznie termiczne i mogłoby wtedy przenosić informację o stanie początkowym, choć oczywiście w przypadku spalonej kartki papieru nie byłby to stan faktycznie dostępny. Niektórzy fizycy uważają z kolei, że oprócz promieniowania pozostałoby również niewielkie ciało mniej więcej o długości i masie Plancka, które przenosiłoby całą informację o stanie początkowym. Z uwagi jednak na to, że istnienie takich obiektów wiąże się z innymi problemami, hipoteza ta nie cieszy się dużą popularnością.

Powstały już setki prac na ten temat, a widoków na rozwiązanie problemu nadal nie ma. W rzeczywistości problem informacyjny ma dwie składowe o różnych stopniach trudności. W przypadku dużych czarnych dziur, czyli dziur, których masa jest znacznie większa od masy Plancka, rozpatrzenie tego problemu powinno być możliwe również bez pełnej wiedzy o grawitacji kwantowej. Tak rzeczywiście jest. Można dyskutować o modelach, w których początkowy, wyłączny stan kwantowy (funkcji falowej) może w każdej chwili przeobrazić się w stan splątania między dziurą a promieniem Hawkinga. Cieplny charakter promieniowania wynika dopiero z pominięcia niektórych subtelnych korelacji w stanie splątanym. Usprawiedliwienie tego pominięcia staje się jasne w momencie uwzględnienia oddziaływania z „otoczeniem”, czyli z kolejnymi stopniami swobody. Podobnie jak inne ciała makroskopowe czarna dziura narażona jest na wpływ otoczenia, a zatem nie można jej rozpatrywać całkowicie odrębnie. W takim wypadku zachodzi proces dekoherencji, który przedstawiliśmy w części 2 tej książki. Informacja o nietermicznym charakterze promieniowania zostaje przekazana w ramach korelacji z otoczeniem, przez co nie może być odczytana lokalnie, czyli na podstawie samego promieniowania. W przypadku czarnych dziur nie dochodzi więc do utraty informacji.

Jeśli czarna dziura straciła już tyle energii związanej z emisją promieniowania, że jej masa zbliża się do masy Plancka, to wtedy za rozwiązanie problemu informacyjnego odpowiada wyłącznie grawitacja kwantowa. To ta bardziej skomplikowana składowa tego problemu.

Grawitacja kwantowa prognozuje wtedy, czy dojdzie do utraty informacji. Teoria ta nie jest jednak w pełni ukończona, a zatem wszelkie spekulacje w tym temacie trzeba uznać za daremne. Przedstawione wyżej ujęcia skwantyzowanej teorii względności i teorii strun nie mogą wprawdzie formułować żadnych stwierdzeń na temat końcowej fazy parowania, ale ich formalizm jest skonstruowany w ten sposób, że nie należałoby oczekiwać żadnej utraty informacji. Zresztą to właśnie postęp w pracy nad teorią strun przyczynił się do tego, że w 2004 roku Hawking zmienił swoje pierwotne stanowisko i poparł hipotezę braku utraty informacji.

W tym samym miejscu, w którym Hawking w 2004 roku wygłosił swój słynny wykład, zagościł inny, jeszcze bardziej znany fizyk. Był to Erwin Schrödinger. Jego wykład w styczniu 1947 roku, na którym obecny był nawet premier Irlandii, spotkał się z równie dużym zainteresowaniem. Autor równania Schrödingera wykorzystał tę okazję, aby zaprezentować swoją zunifikowaną teorię pola, nad którą niezależnie pracował również Einstein. Mimo wielu przedwczesnych pochwał Schrödinger nie sprostał temu zadaniu, podobnie jak Einstein. Na dalsze losy paradoksu informacyjnego możemy czekać z niecierpliwością.

Pierwotne czarne dziury

Czarne dziury emitują promieniowanie. Czy można je zobaczyć? Z uwagi na to, że temperatura Hawkinga jest odwrotnie proporcjonalna do masy czarnej dziury, masa powinna być możliwie niska, ponieważ tylko wtedy dziura będzie dość gorąca, aby można było ją zobaczyć. Czarne dziury, powstałe w wyniku kolapsu grawitacyjnego masywnych gwiazd, mają jednak masę przynajmniej trzykrotnie większą od masy Słońca. Jeśli umieścimy to ograniczenie we wzorze Hawkinga, to otrzymamy temperaturę rzędu jednej stumilionowej części kelwina lub mniejszą. Nie da się tego zobaczyć. Nawet słabe kosmiczne promieniowanie tła, które przenika wszechświat, jest ze swoimi trzema kelwinami dużo „gorętsze”. Czarna dziura pochłania dużo więcej promieniowania tła, niż jest w stanie przekazać w ramach swojego promieniowania.

Zobaczenie promieniowania Hawkinga byłoby możliwe, gdyby istniały dodatkowe wymiary czasoprzestrzeni, o których pisaliśmy w przedostatnim rozdziale. Wtedy w akceleratorze LHC moglibyśmy stworzyć czarne dziury o niższej masie, które byłyby wystarczająco gorące, aby natychmiast wyparować. Po 2008 roku, gdy poznamy dane pochodzące z LHC, przekonamy się, czy realizacja tego przedsięwzięcia jest faktycznie możliwa.

Poza LHC możliwość obserwacji istniałaby właściwie tylko wtedy, gdyby we wczesnej fazie rozwoju wszechświata doszło do powstania lekkich czarnych dziur. Określa się je mianem *pierwotnych* czarnych dziur. Twory takie mogą powstać w procesie kolapsu spontanicznego w wyniku fluktuacji gęstości, które są na tyle duże, aby doprowadzić do dominacji grawitacji nad innymi oddziaływaniami. Wiemy, że wszechświat również we wczesnej fazie rozwoju był dość jednorodny, a fluktuacje temperatury w kosmicznym promieniowaniu tła, uwolnionym 380 000 lat po wielkim wybuchu, sięgają zaledwie milionowej części stopnia. Fluktuacje rozkładają się jednak równomiernie. Ich wartość średnia jest wprawdzie bardzo niska, lecz rozkład w formie krzywej Gaussa w kształcie dzwonu przebiega wokół tej wartości średniej w taki sposób, że nadal istnieje minimalne prawdopodobieństwo wystąpienia wysokiej wartości. Gdyby wszechświat we wczesnej fazie rozwoju był bardzo niejednorodny, takich pierwotnych czarnych dziur byłoby tak wiele, że również obecnie nie mielibyśmy problemu z ich dostrzeżeniem, co rzecz jasna nie odpowiada prawdzie.

Zgodnie z tym, co stwierdziliśmy w części 4 tej książki, powszechnie wychodzi się z założenia, że nasz wszechświat krótko po wielkim wybuchu przeszedł fazę przyspieszonej ekspansji, czyli inflacji. Wzrost ten był tak szybki i intensywny, że wszelkie ewentualne niejednorodności zostały wyprostowane. Dotyczy to także pierwotnych czarnych dziur. Możemy zatem, jeśli w ogóle, obserwować wyłącznie dziury powstałe po okresie inflacji. Obliczenia pokazują, że minimalna masa takich dziur wynosi około jednego grama. Tak lekka dziura wyparowałaby jednak już dawno.

Można więc dowieść, że dziura o masie około 10^{15} grama, powstała we wczesnej fazie rozwoju wszechświata, osiągnęła w obecnym wszechświecie ostatnie stadium odparowania. W stadium tym temperatura gwałtownie rośnie, co sprawia, że można ją zobaczyć. Masa ta odpowiada mniej więcej masie niewielkiej asteroidy. Promień takiej niewielkiej dziury jest jednak mikroskopijny i odpowiada mniej więcej wielkości jądra atomu. Informacje te przypominają tylko o tym, jak bardzo skomplikowane są czarne dziury.

Co pozostałoby po wyparowaniu takiej niewielkiej dziury? Z uwagi na to, że promieniowanie Hawkinga jest zawsze termiczne, obraz byłby zawsze taki sam. W ostatniej fazie odparowywania emitowane są przede wszystkim wysokoenergetyczne promienie gamma o widmie termicznym. Można zatem wypatrywać bezpośrednio takiego zdarzenia bądź poszukiwać pozostałości promieniowania wyemitowanego w nieodległej przeszłości. Mimo licznych prób nie udało się jak dotąd odnaleźć śladów świadczących o istnieniu takiej niewielkiej odparowanej dziury, jednak poszukiwania wciąż trwają.

6. Funkcja falowa wszechświata

Byt to... To nic przeszłego ani przyszłego, bo to coś teraźniejszego, jednostka i wspólnota zarazem.

Parmenides z Elei

Grawitacja kwantowa i kosmologia

Układów kwantowych zasadniczo nie można rozpatrywać indywidualnie. Zgodnie z tym, co stwierdziliśmy w części 2 tej książki, muszą one oddziaływać z innymi układami kwantowymi, które w pewnym sensie pełnią funkcję otoczenia. Następstwem tego jest skrzyżowanie układu kwantowego z jego otoczeniem, które może prowadzić do dekoherencji, czyli klasycznego zachowania układu kwantowego. Dekoherencji można uniknąć wyłącznie wtedy, gdy układ nie posiada zbyt wielu stopni swobody. Tylko w takim przypadku na jaw wychodzi jego prawdziwa natura kwantowa wraz z wszystkimi jej niezwykłymi konsekwencjami. Z uwagi na to, że otoczenie układu posiada z kolei swoje własne otoczenie itd., powstaje cała seria skrzyżowań, która kończy się dopiero na wszechświecie jako całości. Wszechświat jest jedynym naprawdę zamkniętym układem – z definicji nie posiada on żadnego innego otoczenia. Przy założeniu uniwersalnego obowiązywania teorii kwantowej, a jak dotąd nie ma oznak temu przeczących, wszechświat jako całość należy opisywać za pomocą wysoce splątanej funkcji falowej – funkcji falowej wszechświata. Wszechświat staje się kosmosem kwantowym. Podukładom tylko wtedy przypisuje się prawie wyłączną funkcję falową, gdy rozpatruje się je indywidualnie. Przykładem tego mogą być niewielkie układy, takie jak elektrony bądź atomy.

Ze względu na to, że wszechświat obejmuje również wszystkich obserwatorów, aby zachować spójność, konieczne jest włączenie ich do

opisu kwantowego. Kosmologia kwantowa ma zatem bezpośredni wpływ na interpretację zwykłej teorii kwantowej. Amerykańscy fizycy Murray Gell-Mann oraz James Hartle twierdzą nawet, że opierając się na kosmologii kwantowej, można lepiej i dogłębniej zrozumieć mechanikę kwantową.

Argumentacja ta nie ma jeszcze nic wspólnego z grawitacją kwantową. Z uwagi jednak na to, że wszechświat jest w ogromnym stopniu zdominowany przez grawitację, kosmologię kwantową można opisać wyłącznie za pomocą grawitacji kwantowej – analogicznie do opisu kosmologii klasycznej przez ogólną teorię względności Einsteina. W ramach każdego z przedstawionych w poprzedniej części formalizmów jest to możliwe do zrealizowania. Najbardziej intuicyjny jest jednak opis geometrodynamiki kwantowej za pomocą równania Wheelera-DeWitta, dlatego też zasługuje on na najwięcej miejsca. Równanie to dotyczy równania falowego, które stanowi podwaliny równań Einsteina, co oznacza, że właśnie z niego wynikają równania Einsteina w klasycznym przypadku granicznym. Zgodnie z tym, co wyjaśniliśmy w poprzedniej części książki, schemat ten odpowiada postępowaniu Schrödingera podczas poszukiwania funkcji falowej, z której w odpowiednim przypadku granicznym wynika mechanika klasyczna. Równanie Wheelera-DeWitta nie jest zatem wstępnie ograniczone do żadnej określonej skali długości czy energii i powinno również obowiązywać w znacznym oddaleniu od skali Plancka. W przypadku nie nazbyt małego wszechświata niełatwo będzie jednak ukazać różnicę względem klasycznej teorii Einsteina. Uwaga naukowców skupia się więc na poszukiwaniu efektów kwantowych, wywodzących się z wczesnego wszechświata, które pozostają widoczne w ramach obecnych obserwacji kosmologicznych. Do możliwych bezpośrednich efektów kwantowych w kontekście dużego wszechświata jeszcze wrócimy.

W poprzedniej części książki przyjrzeliśmy się bliżej ogólnemu równaniu Wheelera-DeWitta. Stwierdziliśmy przy tym, że funkcja falowa jest określona na bardzo skomplikowanej przestrzeni, którą zgodnie

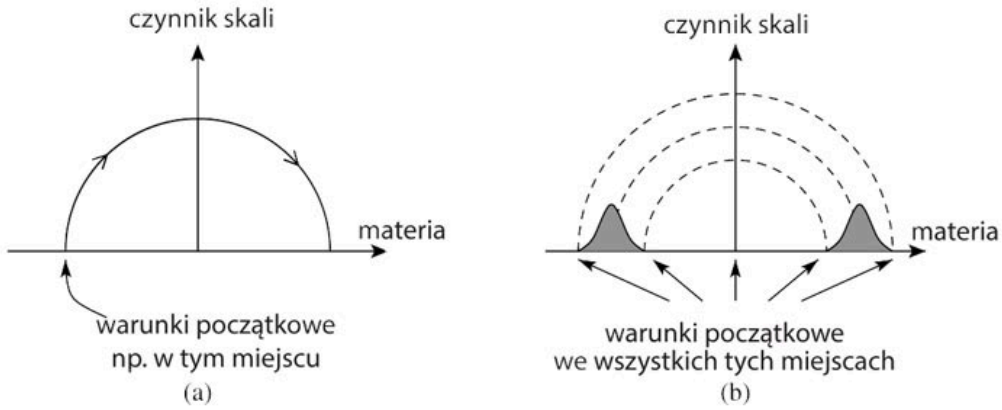
z propozycją Johna Wheelera nazywa się superprzestrzenią. Jest to przestrzeń o nieskończenie wielu wymiarach, w której każdy punkt jest trójwymiarową przestrzenią (a dokładniej: metryką na tej przestrzeni). W ramach kosmologii nasza uwaga nie skupia się jednak na najogólniejszym przypadku. Już w przypadku klasycznym (część 4) doszliśmy do wniosku, że za realistyczne należy uznać rozpatrywanie wielkopowierzchniowych struktur jako prawie jednorodnych i izotropowych. To samo możemy uczynić w przypadku kosmologii kwantowej. Ponadto można przyjąć, że funkcja falowa zależy wyłącznie od czynnika skali („promienia") wszechświata i jednorodnego stopnia swobody materii. W przeciwieństwie do superprzestrzeni o nieskończenie wielu wymiarach przestrzeń konfiguracyjna ma tylko dwa wymiary, a wyznaczają ją czynnik skali oraz materia reprezentowana zazwyczaj przez pole jednorodne. Tę prostą przestrzeń określa się dość dziwnym mianem „minisuperprzestrzeni". W przypadku rozpatrywania modeli, które są trochę bardziej złożone, ale nie osiągają stopnia skomplikowania pełnej teorii (na przykład skwantyzowane czarne dziury), mówi się o „midisuperprzestrzeni" – najwyraźniej nazwy nawiązują do pewnych trendów panujących w modzie w latach sześćdziesiątych XX wieku.

Teraz skoncentrujemy się jednak na minisuperprzestrzeniach. Właściwie modele takie wcale nie są spójne. Powodem tego jest zasada nieoznaczoności. W przypadku nieuwzględnienia stopni swobody (a wskutek założenia jednorodności ignoruje się nieskończenie wiele niejednorodnych stopni swobody) pomija się również przynależne pędy. Stoi to w sprzeczności z zasadą nieoznaczoności Heisenberga, zgodnie z którą położeń (bądź tak jak w tym przypadku uogólnionych położeń) oraz pędów nie można jednocześnie dokładnie określić, a zatem nie mogą one także jednocześnie przyjąć wartości zerowej. Właśnie ten ostatni przypadek tutaj zachodzi. Modele działają więc tylko wtedy, gdy zaniedbanie to nie ma wpływu na istotne wyniki. W rzeczywistości, w ramach zwykłej mechaniki kwantowej, znanych jest wiele sytuacji, w których można bez trudu zrezygnować ze stopni swobody. Niekiedy wystarczy upodobnić jedną cząsteczkę do ciała sztywnego oraz

zignorować stan oscylacji, aby uniknąć sprzeczności związanych z wynikami doświadczalnymi. Funkcja falowa wszechświata jest zatem przede wszystkim określona na minisuperprzestrzeni dwuwymiarowej.

Bryce DeWitt był pierwszym fizykiem, który uwzględnił minisuperprzestrzenie w ramach swojej pracy na temat grawitacji kwantowej, wydanej w 1967 roku. Wtedy nie posługiwał się on jednak określeniem „minisuperprzestrzeń". Nazwę tę zawdzięczamy Charlesowi W. Misnerowi, który w kolejnych latach poświecił temu tematowi wiele czasu. Misner rozpatrywał minisuperprzestrzenie opisujące wszechświat, który był wprawdzie jeszcze jednorodny, ale już nie izotropowy. O minisuperprzestrzeniach mówi się często jako o prostych modelach próbnych grawitacji kwantowej, choć mogą one również służyć do ogólnego opisu prawdziwego kosmosu kwantowego.

W części 5 tej książki wyjaśniliśmy szczegółowo, że fundamentalny świat grawitacji kwantowej jest bezczasowy. Zwykły parametr czasu istnieje wyłącznie w ograniczonych sytuacjach zwanych pół- lub semiklasycznymi. To samo odnosi się oczywiście do kosmologii kwantowej. Funkcja falowa wszechświata (w najprostszym modelu) zależy wyłącznie od czynnika skali i pola materii, a nie od parametru czasu *t*. Stąd biorą się istotne różnice względem kosmologii klasycznej, którą opisaliśmy w części 4. Przyjrzyjmy się na przykład modelowi Friedmana-Lemaître'a, w którym wszechświat rozrasta się, osiąga maksimum, a następnie zaczyna się zapadać. Koncepcja ta nie odpowiada wprawdzie współczesnemu obrazowi stale rozrastającego się wszechświata, idealnie nadaje się jednak do pokazania problemu czasu.



6.1. Grawitacja klasyczna i kwantowa odznaczają się nader różnymi pojęciami czasu

Na rysunku 6.1 w części (a) przedstawiono ewolucję klasyczną. Odpowiada ona istnieniu klasycznego toru. (Pod pojęciem toru rozumiemy tutaj krzywą ewolucji wszechświata w jego przestrzeni konfiguracyjnej). Warunki początkowe można określić na przykład na jednym końcu toru (zazwyczaj w przypadku niewielkiego wszechświata), gdzie czas rozpoczyna swój bieg (tor ruchu). Zapadająca się część krzywej jest deterministycznym odpowiednikiem części rozrastającej się. Kierunek jest w tym wypadku dowolny – warunki początkowe można ustalać na końcu zlokalizowanym zarówno z lewej (zgodnie z tym, co przedstawiono w części (a)), jak i prawej strony. Decydujące jest samo istnienie toru.

W wersji kwantowej (rys. 6.1b) klasycznego toru nie ma. Występuje tylko funkcja falowa, która jest rozwiązaniem równania Wheelera-DeWitta. W sytuacjach zbliżonych do klasycznych najlepszym rozwiązaniem byłoby stworzenie funkcji falowej, która niczym wąż ogrodowy oplatałaby klasyczny tor. Z uwagi jednak na to, że tor nie istnieje, określenie warunków początkowych na jednym z jego końców również nie jest możliwe. Sformułowanie warunków początkowych jest konieczne w przypadku funkcji falowej. Struktura tego równania (które jest równaniem falowym) sugeruje, aby określić warunki początkowe wzdłuż całej linii widocznej na powyższym diagramie, czyli wyznaczyć

funkcję falową przy niewielkim czynniku skali dla wszystkich wartości materii. Z uwagi na to, że tor nie istnieje, rozmowa o wielkim wybuchu i wielkim krachu nie ma sensu – można jedynie dokonać rozróżnienia między małym a dużym wszechświatem. W celu otrzymania węża falowego, który przynajmniej w przybliżeniu przypominałby klasyczny tor pokazany w części (a), należy zarówno z lewej, jak i prawej strony wyznaczyć „pakiet falowy", służący za warunek początkowy. W przeciwnym razie otrzymane rozwiązanie będzie rozciągnięte z każdej strony i nie będzie w niczym przypominać klasycznego toru z części (a). W kosmologii kwantowej klasycznie zapadająca się część istnieje zatem od samego początku. Odpowiada to obrazowi „czasu wewnętrznego", który przedstawiliśmy w części 5. Czas wewnętrzny jest w tym przypadku czynnikiem skali, a stary czas *t* znika wraz z klasycznym torem. Sytuacja zobrazowana na rysunku 6.1 pozostanie nadal istotna w odniesieniu do omawianej poniżej strzałki czasu.

Czego oczekuje się od kosmologii kwantowej? Jaką informację skrywa funkcja falowa wszechświata? Oczywiście najważniejszym zadaniem pozostaje rozwiązanie niewyjaśnionych jeszcze problemów kosmologii klasycznej (część 4) – jednym z najistotniejszych jest problem osobliwości. Powstania wszechświata nie da się opisać przy wykorzystaniu ogólnej teorii względności, a „wielki wybuch" to obszar nieskończenie wielkiej krzywizny czasoprzestrzeni i gęstości energii materii. Czy kosmologia kwantowa może rozwiązać problem osobliwości? Już DeWitt w 1967 roku sugerował, że w przypadku granicznym znikającego czynnika skali funkcja falowa musiałaby być równa zeru. W takim przypadku zgodnie z mechaniką kwantową klasyczny wszechświat nie miałby prawa istnieć w tym miejscu, co oznaczałoby brak wielkiego wybuchu i klasycznej osobliwości. W ramach równania Wheelera-DeWitta pozostaje kwestią otwartą, czy ten brak osobliwości jest nieodwołalny, czy też nie. Istnieją modele i rozwiązania, w których tak jest, oraz takie, w których tak nie jest. Te ostatnie są jednak nieprzydatne, ponieważ uniwersalna teoria fizyczna nie może obejmować żadnych osobliwości.

Poniżej omówimy kosmologię kwantową w ramach grawitacji pętlowej, w której uniknięcie osobliwości jest łatwiejszym zadaniem.

W przypadku osobliwości kosmologicznych automatycznie na myśl przychodzi wielki wybuch i wielki krach. Istnieją modele, w ramach których osobliwości mogą występować również w dużym wszechświecie, ale nie ma pewności co do tego, czy faktycznie opisują one nasz wszechświat. Jest to jednak możliwe. Z fenomenem ciemnej energii mieliśmy już do czynienia w części 4 tej książki. W najprostszym przypadku fenomen ten wiązał się ze stałą kosmologiczną, a w najbardziej skomplikowanych modelach – z zależną w czasie energią próżni, wywołaną przez pole materii. W przypadkach tych wszechświat może w przyszłości podążać na przykład w kierunku „wielkiego rozdarcia", o którym wspomnieliśmy w części 4. Zgodnie z tą koncepcją wszechświat w skończonym czasie rozrasta się w sposób nieskończony, a efektem tego jest ostatecznie jego rozdarcie. Inna osobliwość dotycząca przyszłości występuje w przypadku obecności niezwykłych pól o szczególnych właściwościach. Na podstawie poczynionych do tej pory obserwacji nie można jednak kategorycznie wykluczyć ich istnienia. Gdyby założenia te okazały się prawdziwe, w przyszłość wszechświata byłaby wpisana inna osobliwość nazywana „wielkim zatrzymaniem". W sytuacji takiej wszechświat redukuje prędkość ekspansji w tak szybkim tempie, że natychmiast dochodzi do jego zatrzymania, a przyspieszenie ujemne (hamowanie) staje się nieskończenie wielkie, czyli osobliwe.

Co na to kosmologia kwantowa? Podczas poszukiwania rozwiązania równania Wheelera-DeWitta o rozsądnych właściwościach matematycznych okazuje się, że wszystkie te rozwiązania sprowadzają się do funkcji falowej, która znika w miejscu wystąpienia osobliwości, co w całości pokrywa się z rozważaniami DeWitta. Prawdopodobieństwo wystąpienia osobliwości jest zatem zerowe, a tym samym pomijalne w kosmologii kwantowej, co jest wynikiem zadowalającym. W modelu tym to samo odnosi się do wielkiego wybuchu, co oznacza, że również on

nie występuje w teorii kwantowej. Osobliwości są więc artefaktami teorii klasycznej.

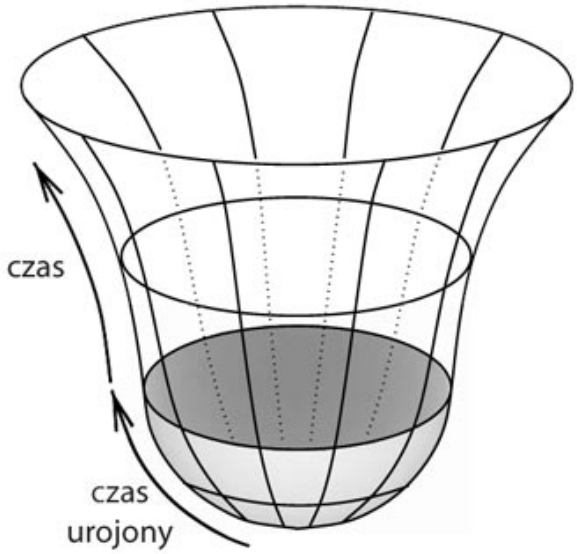
Z problemem unikania osobliwości ściśle powiązany jest problem warunków brzegowych i początkowych w kosmologii kwantowej. Kwestia ta była już istotna w odniesieniu do rysunku 6.1. Kto wyznacza te warunki początkowe? W ramach zwykłej fizyki kwantowej tym kimś jest osoba prowadząca eksperyment w laboratorium, która przygotowuje pożądany stan kwantowy. We wczesnym okresie rozwoju wszechświata osoba taka jednak nie istniała. Stwierdzenie to dało początek rozważaniom, że być może dogłębne zrozumienie pełnej teorii i jej struktur matematycznych ukaże sensowność istnienia tylko jednego rozsądnego rozwiązania (lub umiarkowanej wielkości klasy rozwiązań). W pewnym sensie doszłoby zatem do zbiegnięcia się warunków brzegowych z dynamiką, a co za tym idzie – do obalenia wprowadzonego przez Newtona i wspieranego przez Wignera podziału opisu przyrody na prawa dynamiczne i dowolne warunki początkowe.

Autorami propozycji, która wysuwała takie postulaty, są fizycy James Hartle oraz Stephen Hawking, a nazywa się ją „warunkiem braku granic" (zgodnie z nieco koślawym tłumaczeniem oryginalnego *no-boundary condition*). Co to znaczy? W części 5 omówiliśmy metodę całki po trajektoriach Feynmana. W odniesieniu do grawitacji kwantowej oznacza to, że trzeba zsumować wszystkie możliwe czasoprzestrzenie wraz z ich metrykami, aby otrzymać pożądaną funkcję falową. Zrozumienie „warunku braku granic" wymaga uwzględnienia kolejnych dwóch aspektów. Po pierwsze, Hartle i Hawking rozpatrują tak zwaną euklidesową całkę po trajektoriach. Zakłada ona sumowanie po czterowymiarowych przestrzeniach, a nie (jak można by oczekiwać) po czterowymiarowych czasoprzestrzeniach, czego efektem jest wymazanie różnicy między przestrzenią a czasem. W formalizmie rozwiązuje się to poprzez wprowadzenie „czasu urojonego", co sprowadza się do tego, że zwykły czas staje się czwartą współrzędną przestrzenną. Hartle i Hawking

wychodzą z założenia, że suma po przestrzeniach jest równoważna sumie po czasoprzestrzeniach, i w ogóle jest pierwszą jednoznaczną sumą.

Drugi aspekt odnosi się do znaczenia „braku granic". Hartle i Hawking proponowali ogólną zasadę, według której czterowymiarowe przestrzenie będące przedmiotem sumowania, podobnie jak powierzchnia kuli, nie posiadają żadnej trójwymiarowej krawędzi. (W rzeczywistości istnieje oczywiście jedna krawędź trójwymiarowa, którą jest współczesny nam wszechświat; w domyśle chodzi zatem o brak innej krawędzi w „przeszłości"). Według Hartle'a i Hawkinga dzięki temu możliwe powinno być uniknięcie osobliwości wielkiego wybuchu, ponieważ w przestrzeniach objętych całkowaniem po trajektoriach nie ma żadnej krawędzi, a tym samym żadnej granicy. Naukowcy zastanawiają się również nad tym, czy warunek ten nie mógłby doprowadzić do sformułowania jednoznacznej funkcji falowej wszechświata. Dodatkowe warunki brzegowe byłyby wtedy nadmiarowe, a podział na dynamikę i warunki brzegowe – niepotrzebny.

Proponować zasadę to jedno, a przewidzieć związane z tym konsekwencje to drugie. Dokładne obliczenie całki po trajektoriach jest w zasadzie niemożliwe, często jednak wystarczy ustalenie, jakie geometrie są dominujące w sumie całek po trajektoriach. Dzięki temu można obliczyć szukaną funkcję falową. Hartle i Hawking stwierdzili, że w kosmologii kwantowej, uwzględniającej stałą kosmologiczną, dominuje geometria przedstawiona na rysunku 6.2, nazywana często „instantonem Hartle'a-Hawkinga".



6.2. Instanton Hartle'a-Hawkinga

Co właściwie widzimy? Rysunek ten ukazuje geometrię (to znaczy jej dwuwymiarowy przekrój), w której dolna połowa (w rzeczywistości czterowymiarowej) powierzchni kuli jest przyklejona do górnej połowy przestrzeni de Sittera, zobrazowanej na rysunku 4.10. Na dole nie ma zatem żadnej krawędzi. Czy w rzeczywistości górną częścią nie jest jednak czasoprzestrzeń? Tak właśnie jest. Formalizm matematyczny płata czasem figle – w tym wypadku w celu zachowania spójności wymusza on ponowną transformację górnej części powierzchni kuli (której spodziewalibyśmy się zgodnie z formalizmem euklidesowym) w czasoprzestrzeń. W rzeczywistości jest to nawet efekt pożądany. Hartle i Hawking interpretują ten stan rzeczy jako powstanie zwykłego czasu z bezczasowego świata samej przestrzeni. Czas skończyłby się więc (z naszej perspektywy patrząc wstecz) na długo przed „wielkim wybuchem". Zaprezentowany na rysunku 6.2 instanton jest jednak zaledwie przyczynkiem, choć głównym, do całki po trajektoriach. Bezpośrednie znaczenie fizyczne ma natomiast wynik całkowania po trajektoriach, czyli funkcja falowa.

Dość szybko okazało się, że propozycja Hartle'a i Hawkinga wbrew pierwotnym oczekiwaniom nie prowadzi do sformułowania jednoznacznej funkcji falowej wszechświata. W związku z tym należy zatem dokonać kolejnego wyboru i rozwiązać problem warunków brzegowych. „Warunek braku granic" dostarcza nam przynajmniej ograniczoną klasę funkcji falowych, którą można przeanalizować. Ponadto zdaje się on prognozować, że we wczesnym okresie rozwoju wszechświata niegrawitacyjne stopnie swobody znajdowały się w stanie podstawowym (stan próżni). Stan próżni, przyjęty na potrzeby omawianych w części 4 fluktuacji kwantowych, mógłby być więc uzasadniony w obliczu akceptacji warunku Hartle'a i Hawkinga w ramach kosmologii kwantowej.

Oczywiście istnieją również inne propozycje. Podobną popularnością jak „warunek braku granic" cieszyła się propozycja Aleksandra Wilenkina z 1982 roku, znana pod nazwą „warunek tunelowy". Opisany przez mechanikę kwantową efekt tunelowy omówiliśmy krótko w części 2. W ramach tego zjawiska cząstki kwantowe mogą przezwyciężyć barierę energetyczną, do czego nie są zdolne w przypadku klasycznym. Za przykład może posłużyć rozpad alfa, w którym cząstka alfa (składająca się z dwóch protonów i dwóch neutronów) zostaje wyemitowana przez jądro atomu. Ze względu na panujące w jądrze warunki energetyczne cząstka alfa nigdy nie opuściłaby jądra jako cząstka klasyczna, ponieważ nie posiada energii potrzebnej do pokonania bariery energetycznej. Jako cząstka kwantowa o właściwościach falowych może ona jednak wydostać się na zewnątrz dzięki zjawisku tunelowemu. Dla równania Wheelera-DeWitta Wilenkin wybrał więc rozwiązanie, które spełnia wymogi formalne funkcji falowej wyemitowanej cząstki alfa. Sam Wilenkin opisuje całe zjawisko nieco metaforycznie, mówiąc, że wszechświat pojawił się w jego życiu z niczego dzięki efektowi tunelowemu.

Zgodnie z powyższym przekonaliśmy się zatem, że w grawitacji kwantowej nie ma ogólnego parametru czasu. Co to znaczy, gdy mówimy, że wszechświat dzięki efektowi tunelowemu pojawia się z niczego? Cząstka alfa pojawia się dzięki efektowi tunelowemu w odniesieniu do

zwykłego parametru czasu *t*: po uwięzieniu w jądrze istnieje pewne prawdopodobieństwo, że dojdzie do jej ucieczki w przyszłości. Rozumowanie to w przypadku wszechświata nie jest jednak możliwe, ponieważ funkcja falowa Wilenkina jest statyczna – posiada wyłącznie *formę* funkcji falowej cząstki alfa. W związku z tym wykorzystanie efektu tunelowego jest w tym wypadku nieco przesadzone.

Interesująco brzmią oczywiście możliwe prognozy, które odróżniają rozwiązanie Wilenkina od koncepcji Hartle'a i Hawkinga. Można na przykład pochylić się nad kwestią prawdopodobieństwa istnienia inflacyjnej fazy wszechświata (część 4). W tym temacie oba rozwiązania zdają się dochodzić do różnych wniosków. W przypadku warunku tunelowego inflacja jest w wysokim stopniu prawdopodobna. „Warunek braku granic" dowodzi natomiast czegoś zupełnie odwrotnego. Ostatnie słowo w tej kwestii nie zostało jeszcze powiedziane. Jeśli jednak pragniemy pozostać przy potwierdzonym empirycznie modelu inflacyjnego wszechświata, to o „warunku braku granic" trzeba raczej zapomnieć.

W części 5 pisaliśmy o sytuacjach, w których zwykły czas wyłania się z bezczasowej grawitacji kwantowej w formie przybliżenia. Przykładem analogicznym jest przypadek graniczny optyki geometrycznej, wywodzący się z fundamentalnej optyki falowej. Podobnie jak w tym przypadku promienie świetlne powstają z fal elektromagnetycznych, tak w zbliżonych okolicznościach odnaleźć można czasoprzestrzenie w bezczasowej grawitacji kwantowej („tory przestrzeni w czasie"). To samo odnosi się również do kosmologii kwantowej. Z realistycznego punktu widzenia należałoby jednak uzupełnić oba jednorodne stopnie swobody (czynnik skali i pole materii) niejednorodnymi stopniami swobody. Nasz wszechświat nie jest przecież całkowicie jednorodny, bo w przeciwnym razie nie byłoby nas. Dużego przybliżenia należy się spodziewać, gdy te niejednorodne stopnie swobody będzie można rozpatrywać jako niewielkie perturbacje na jednorodnym tle. Pod względem fizycznym są to na przykład niewielkie fluktuacje gęstości, z których mogą powstać galaktyki

(część 5), lub słabe fale grawitacyjne, które przenikają cały wszechświat. Następnie przy przyjętym przybliżeniu jednorodne stopnie swobody *określają* semiklasyczny parametr czasu *t*, a niejednorodne zmienne jako stany kwantowe ewoluują w odniesieniu do tego *t* – wszystko odbywa się zgodnie z równaniem Schrödingera. W rezultacie nie otrzymujemy jednak kierunku upływu czasu – do tego brakuje nam decydującej informacji na temat entropii (zob. poniżej). Jeśli rzeczywiście dopiero model Friedmana-Lemaître'a, który opisuje jednorodne stopnie swobody, określa czas, to ekspansja wszechświata staje się tautologią – wszechświat nie rozszerza się w czasie, lecz dopiero definiuje czas. Zegary fizyczne, opisane przez niejednorodne stopnie swobody, nie są tautologiczne, lecz logiczne, ponieważ są powiązane ze zmiennymi określającymi czas (czynnik skali i pole).

Temat trafności zastosowania powyższego przybliżenia parametru czasu w kosmologii kwantowej rozważymy jeszcze w następnym rozdziale. Rozwiązania tego z pewnością nie można jednak stosować w każdym wypadku. Z równania Wheelera-DeWitta wynika, że zawsze istnieją regiony, w których funkcja falowa na wielu obszarach będzie rozmazana, czyli nie wszędzie wąż będzie ciasno oplatał klasyczny tor. Na poziomie fundamentalnym kosmos kwantowy jest bezczasowy.

W części 5 oprócz geometrodynamiki omówiliśmy również bardziej nowoczesną wersję posługującą się pętlami. Główną konkluzją była wtedy dyskretna natura przestrzeni w najmniejszych skalach. Jaki wpływ ma to stwierdzenie na kosmologię kwantową? Jak dotąd z pełnej teorii pętli nie udało się dokładnie wyodrębnić kosmologicznego przypadku granicznego. Rozsądnym posunięciem wydaje się jednak bezpośrednie przeniesienie dyskretnej struktury przestrzeni na czynnik skali, czyli wielkość wszechświata. Dzięki temu docieramy bezpośrednio do tak zwanej pętlowej kosmologii kwantowej. Za twórcę tej teorii uznaje się pracującego obecnie na Pennsylvania State University fizyka Martina Bojowalda, który ogłosił ją w 1999 roku.

Dyskretna natura przestrzeni prowadzi do zastąpienia równania Wheelera-DeWitta równaniem rekurencyjnym. Jest to równanie, w którym dozwolone są wyłącznie określone dyskretne wartości czynnika skali. Informuje ono zatem, w jaki sposób w skończonej liczbie kroków przeskoczyć od jednej do drugiej dozwolonej wartości. Ta ziarnista natura daje o sobie znać dopiero w przypadku niewielkiego wszechświata. W skalach, które przekraczają kilka długości Plancka, nie ma ona większego znaczenia. W takich wypadkach nie ma różnicy między równaniem rekurencyjnym a równaniem Wheelera-DeWitta.

Kosmologia pętlowa w dość przekonujący sposób unika klasycznej osobliwości: wśród dyskretnych wartości czynnika skali brakuje wartości zerowej, odpowiadającej osobliwości, czego konsekwencją jest brak „wielkiego wybuchu”! Nieco osobliwy jest jednak fakt, że wartość zerową można wprawdzie przeskoczyć, ale powoduje to przejście do „innego wszechświata”, który w geometrodynamice w ogóle nie występuje. Znaczenie fizyczne tego nowego obszaru czeka jeszcze na wyjaśnienie. Kolejną cechą kosmologii pętlowej jest prognozowana przez nią modyfikacja starego modelu Friedmana-Lemaître’a przy odpowiednio niewielkim wszechświecie. Zmiana ta zdaje się prowadzić do automatycznego wywołania inflacyjnej fazy wszechświata bez konieczności angażowania nieznanych pól pochodzących z teorii zunifikowanych.

Kosmologia kwantowa wydaje się posiadać wszelkie atuty potrzebne do rozwiązania niewyjaśnionych problemów pojęciowych kosmologii klasycznej. Czy istnieją jednak bezpośrednie badania doświadczalne? Na to pytanie nie ma prostej odpowiedzi. Odstępstwa od kosmologii klasycznej zdają się odnosić przede wszystkim do wczesnego okresu rozwoju wszechświata, o którym wiemy dość niewiele. Najstarszym zdarzeniem, znanym nam dzięki odebraniu bezpośredniego sygnału, jest uwolnienie kosmicznego promieniowania tła 380 000 lat po wielkim wybuchu. Rysunek 4.11 (część 4) przedstawia dokładne spektrum tego promieniowania – rozkład gorętszych i chłodniejszych miejsc. Spektrum to

zawiera ogromną ilość informacji. Jak dotąd do interpretacji prowadzonych obserwacji wykorzystywano głównie kosmologię klasyczną (zgodnie z jej opisem w części 4). Kosmologia kwantowa pozostawia ewentualnie swoje ślady w formie niewielkich odstępstw od tego spektrum, których jednak nie możemy dostrzec na podstawie aktualnych danych. Czy przyszłe misje kosmiczne będą w stanie wytropić te ślady, pozostaje kwestią otwartą. Bardziej bezpośrednie wskazówki mogą być ukryte w falach grawitacyjnych wyemitowanych we wczesnym okresie rozwoju wszechświata, ułamki sekundy po wielkim wybuchu (porównaj z częścią 4). Ich istnienie mogą udowodnić planowane misje LISA oraz *Big Bang Explorer*. Także w tym kontekście z niecierpliwością czekamy na to, co przyniesie przyszłość.

Jeden czy więcej światów?

Kosmologia kwantowa to odniesienie teorii kwantowej do wszechświata jako całości. Centralną wielkością, którą poznaliśmy w ostatnim rozdziale, jest tutaj funkcja falowa wszechświata. Z definicji opisuje ona wszystkie istniejące stopnie swobody, czyli również wszystkich obserwatorów. W związku z tym kosmologia kwantowa musi zawsze zawierać spójny opis takiego wewnętrznego obserwatora.

Hipotetyczne spojrzenie z „zewnątrz" ujawniłoby bezczasowy świat kwantowy, w którym brakuje oznak klasycznego świata. Już wcześniej przekonaliśmy się, że spojrzenie wewnątrz może ponownie przywołać aproksymatywne pojęcie czasu: niektóre stopnie swobody (do których zalicza się także wielkość wszechświata) definiują standard czasu, względem którego rozwijają się pozostałe stopnie swobody. W części 5 przytoczyliśmy przykład książki jako całości, która pozostaje w związku z czytelnikiem, ale jednocześnie przedstawia iluzję rozwoju w czasie. Podobnie jest w przypadku właściwości klasycznych.

W części 2 przekonaliśmy się, że układ kwantowy może przyjąć właściwości klasyczne, gdy jest narażony na niekontrolowane i nieuniknione oddziaływanie ze strony otoczenia. Jest to proces

dekoherencji. W przypadku wszechświata jako całości nie można jednak mówić o dekoherencji, ponieważ wszechświat zawiera już wszystko, a zatem nie ma żadnego otoczenia. W związku z tym dekoherencja jest uzasadniona wyłącznie w przypadku części wszechświata (podukłady) w odniesieniu do obserwacji wykonanych przez inne układy częściowe („obserwatorzy"). Pozostanie przy liniowej strukturze teorii kwantowej, czyli uniwersalnym obowiązywaniu zasady superpozycji, oznacza istnienie w funkcji falowej wszechświata miriadów odgałęzień z tymi samymi obserwatorami widzącymi świat o klasycznych właściwościach. Wyłącznie w przypadku doraźnie wprowadzonego „kolapsu funkcji falowej" (część 2) ten świat kwantowy ograniczyłby się do jednego odgałęzienia. Jak dotąd brakuje jednak dowodów empirycznych wskazujących na wystąpienie takiego kolapsu. Wszystkie ujęcia dotyczące grawitacji kwantowej, przedstawione w części 5 tej książki, są liniowe i przynajmniej pośrednio przyjmują teorię kwantową według interpretacji Everetta.

Grawitacja kwantowa dopuszcza przede wszystkim możliwość istnienia stanów kwantowych, które opisują superpozycję różnych metryk, czyli na pozór zróżnicowanych wszechświatów. Jest to swoista analogia do kota Schrödingera. Przed chwilą stwierdziliśmy jednak, że wszechświat jako całość nie posiada żadnego otoczenia. Jak zatem może dojść do dekoherencji? Najpierw trzeba zaznaczyć, że słowo „otoczenie" nie zawsze należy rozumieć dosłownie – w tym wypadku chodzi o nieistotne w szerszym kontekście stopnie swobody, które oddziałują z danym układem, lecz same pozostają niedostępne. W ten sposób w przypadku wszechświata może dojść do powstania otoczenia w formie niewielkich niejednorodności, takich jak fluktuacje gęstości czy fale grawitacyjne. Kosmologicznemu obserwatorowi dostępne będą wtedy wyłącznie wielkości globalne (na przykład czynnik skali) i twory makroskopowe, takie jak gwiazdy, ale nie słabe fale grawitacyjne, które są obecne wszędzie. Te mogą stanowić więc otoczenie w rozumieniu dekoherencji. W rzeczywistości szczegółowe obliczenia pokazują, że oddziaływanie ze słabymi falami grawitacyjnymi wystarcza, aby globalną strukturę wszechświata przedstawić w formie klasycznej. Z uwagi na to, że

grawitacja łączy się w ten sam sposób ze wszystkimi formami energii, jest ona z jednej strony narażona na silny wpływ dekoherencji, a z drugiej strony sama wymusza klasyczny wygląd materii, czyli, mówiąc potocznie, materia zostaje „zmierzona” przez grawitację i odwrotnie.

W przypadku dużego wszechświata dekoherencja jest zazwyczaj do tego stopnia efektywna, że klasyczny obraz przedstawiony w części 4 pozostaje w dużej mierze niezmieniony. W odniesieniu do małego wszechświata sytuacja wygląda inaczej i należy odwołać się do pełnego opisu kwantowego. W rzeczywistości powstaje mniej więcej taki scenariusz: jeśli cofniemy się w czasie w miarę rozwoju naszego wszechświata, to mniej więcej do początku okresu inflacji widoczny jest klasyczny obraz jego ewolucji. Przed tym okresem (czyli wcześniej) dekoherencja jest pomijalna, inne odgałęzienia funkcji falowej wszechświata pokrywają się z naszym odgałęzieniem, a ogólna sytuacja odpowiada istnieniu bezczasowego świata kwantowego. Jeśli liczymy zatem wstecz, począwszy od dnia dzisiejszego, to ewolucja klasyczna zatrzymuje się przed „wielkim wybuchem” i kończy we mgle bez dynamiki. Iluzja czasu ulatnia się.

Interpretacja prawdopodobieństwa teorii kwantowej w odniesieniu do kosmosu kwantowego zdaje się w pierwszej chwili nie mieć sensu, ponieważ wszystkie możliwości zgodne z całą funkcją falową istnieją w rzeczywistości jako odgałęzienia. Prawdopodobieństwa trzeba rozważyć pośrednio jedynie wtedy, gdy rozpatrujemy względną częstość poszczególnych odgałęzień i jednocześnie zakładamy, że żyjemy w całkowicie nietypowym świecie. Potem jednak spójny obraz dostarcza nam przedstawiona w części 2 interpretacja Everetta. Obserwatorzy jako wielkości lokalne są częścią globalnej funkcji falowej. Na podstawie ich obserwacji i z uwagi na brak informacji o innych odgałęzieniach powstaje obraz świata kwantowego, który rozwija się indeterministycznie. W związku z tym powstaje wyobrażenie jedynego świata kwantowego o miriadach klasycznych „światów częściowych”. W jednym z następnych rozdziałów, poświęconym zasadzie antropicznej, powrócimy do obu tych

kwestii. Najpierw zastanowimy się jednak nad uzasadnieniem powstania kierunku upływu czasu w ramach bezczasowego świata kwantowego.

Powstanie kierunku upływu czasu

Rozwój świata w czasie cechuje się nieodwracalnością. W części 3 poznaliśmy drugą zasadę termodynamiki, zgodnie z którą entropia zamkniętego układu, będąca miarą nieporządku, nigdy nie maleje, lecz dąży do osiągnięcia wartości maksymalnej. Do niej w naszym wszechświecie dużo jej jeszcze brakuje, czego powodem jest dość młody wiek naszego wszechświata oraz fakt, że jego rozwój rozpoczął się od bardzo niewielkiej entropii. To różnica między zimną przestrzenią kosmiczną a gorącymi gwiazdami ukazuje, że bardzo daleko nam jeszcze do stanu równowagi, w którym entropia osiąga wartość maksymalną. Do powstania kierunku upływu czasu („strzałki czasu") doszło zatem w nieprawdopodobnym stanie początkowym wszechświata. Przewidział to już Boltzmann. Według niego za wyjaśnienie tego, że w przyrodzie przejście od stanu prawdopodobnego do nieprawdopodobnego nie zachodzi tak często jak w kierunku odwrotnym, powinno wystarczyć założenie bardzo nieprawdopodobnego stanu początkowego całego otaczającego nas wszechświata.

W przypadku zwykłych gazów stanem maksymalnej entropii jest rozkład równomierny, a więc stan makroskopowo jednorodny. W obecności grawitacji jest jednak inaczej! Z uwagi na to, że jest ona powszechną siłą przyciągania, w sposób naturalny dochodzi do tego, iż układ staje się coraz bardziej niejednorodny i „skleja się", tworząc struktury. Temu entropowemu trendowi zawdzięczamy zatem istnienie galaktyk. Druga zasada termodynamiki w przypadku układów grawitacyjnych prowadzi tym samym do zmiany stanu jednorodnego na niejednorodny. Początek wszechświata musiał być bardzo jednorodny. Na podstawie wzoru na entropię czarnych dziur (część 3) można powiedzieć, że entropia naszego wszechświata byłaby maksymalna, gdyby cała materia tkwiła w gigantycznej czarnej dziurze. Byłby to więc stan

maksymalnego prawdopodobieństwa. Dlaczego wszechświat nie rozpoczął się w tak prawdopodobnym stanie?

Ten prawdopodobny stan byłby właściwie wyznaczony przez tak zwaną białą dziurę, czyli przeciwieństwo czarnej dziury. Z tworu takiego obiekty mogłyby wypływać, ale nigdy nie mogłyby do niego wnikać, czyli dokładnie odwrotnie niż w przypadku czarnej dziury. Wszechświat we wczesnej fazie rozwoju był jednak bardzo jednorodny i nie zawierał żadnych białych dziur. Dlaczego? Z uwagi na to, że ogólna teoria względności za początek wszechświata uznaje osobliwość, nie można udzielić odpowiedzi na to pytanie. Mimo to przekonaliśmy się, że kosmologia kwantowa może prowadzić do uniknięcia osobliwości. Czy potrafi ona również wyjaśnić strzałkę czasu?

W tym miejscu można się nieco zdziwić. Czy nie widzieliśmy, że fundamentalny świat kwantowy jest bezczasowy? Jak z nieistniejącego czasu ma powstać kierunek upływu czasu? W rzeczywistości w świecie kwantowym na najbardziej podstawowym poziomie nie ma ani czasu, ani kierunku upływu czasu. Parametr czasu pojawia się dopiero w przybliżeniu semiklasycznym i właśnie wtedy uzasadnione jest pytanie o to, czy grawitacja kwantowa nadaje kierunek temu rzeczywistemu parametrowi.

Decydujące znaczenie w tym względzie ma fakt, że równanie Wheelera-DeWitta jest asymetryczne w odniesieniu do „czasu wewnętrznego" (a w zasadzie do wielkości wszechświata; zob. część 5). W przypadku niewielkiego wszechświata (niski czynnik skali) przyjmuje ono bardzo prostą, a w przypadku dużego wszechświata (wysoki czynnik skali) – bardziej skomplikowaną postać. Asymetria względem czasu wewnętrznego jest w nim w pewnym sensie zaprogramowana. Czas semiklasyczny, jeśli w ogóle się pojawia, jest jednak związany z czasem wewnętrznym, a zatem również aproksymatywny parametr czasu *t* dziedziczy możliwy kierunek upływu czasu.

Kolejnym warunkiem pojawienia się w tym miejscu kierunku upływu czasu jest nieskomplikowana postać funkcji falowej przy niewielkich

czynnikach skali. Taki warunek brzegowy wyraziłby w zasadzie, że w tym obszarze nie występuje żadne skrzyżowanie między czynnikiem skali a innymi stopniami swobody. Warunek ten dopuszcza prosta forma równania Wheelera-DeWitta dla niewielkich wszechświatów. Z uwagi na asymetrię równania w przypadku większego wszechświata wystąpienie splątania jest nieuniknione. Z naszych rozważań z części 3 wiemy jednak, że rosnącemu splątaniu towarzyszy także rosnąca entropia każdego układu będącego częścią splątania, czego przyczyną jest brak informacji o niedostępnych stopniach swobody w obrębie splątania. Entropia jest wtedy powiązana z czynnikiem skali: niewielki wszechświat – niewielka entropia; duży wszechświat – duża entropia. Ekspansja wszechświata określa więc kierunek upływu czasu i tym samym wyznacza upragnioną pierwotną strzałkę czasu, wspólną dla wszystkich strzałek czasu, które omówiliśmy w części 3. Wszechświat nie rozszerza się zatem w czasie, lecz definiuje czas *oraz* jego kierunek.

Scenariusz taki opiera się oczywiście wyłącznie na przypuszczeniach, ale jednocześnie uwidacznia możliwość znalezienia spójnego i zadowalającego rozwiązania problemu nieodwracalności w ramach kosmologii kwantowej. Sformułowanie ostatecznej odpowiedzi musi jednak poczekać na pełne zrozumienie grawitacji kwantowej. Kwestią otwartą pozostaje przede wszystkim pytanie, czy struktura matematyczna równania Wheelera-DeWitta (lub też jej uogólnienie w teorii pętli i strun) pozwala tylko na ograniczoną klasę rozwiązań o pożądanych właściwościach. Odpowiadałoby to pierwotnym nadziejom DeWitta i oznaczałoby – rozważane przez nas w poprzednim rozdziale – zniesienie podziału na warunki brzegowe i dynamikę.

W nawiązaniu do relacji między entropią a wielkością wszechświata nasuwa się zatem pytanie: Co, jeśli wszechświat rozrasta się zgodnie z obrazem klasycznym, osiągnie maksimum, a następnie ponownie zacznie się zapadać, czyli zmierzać do wielkiego krachu? Aktualne obserwacje, mówiące o przyspieszonym wzroście wszechświata, oczywiście nie potwierdzają tego scenariusza, ale pytanie to ma dla nas zasadnicze

znaczenie. Ponadto nie wiadomo, czy rzeczywiście nie dojdzie do wielkiego krachu. Czy w razie zaistnienia takiego modelu należy również oczekiwać, że nastąpi odwrócenie entropii, a tym samym strzałki czasu w punkcie zwrotnym? Czy kawałki rozbitej filiżanki ponownie połączą się w jedną całość, gdy wszechświat zacznie się zapadać?

Z uwagi na to, że funkcja falowa wszechświata zależy wyłącznie od wielkości wszechświata, a nie od czasu zewnętrznego, nie potrafi ona odróżnić wielkiego wybuchu od wielkiego krachu (rys. 6.1). Jeśli zatem entropia rzeczywiście rośnie wraz ze zwiększającą się wielkością, to strzałka czasu zawsze określa ekspansję. Po osiągnięciu klasycznego maksimum nie moglibyśmy jednak zaobserwować procesu naprawy filiżanki. Po pierwsze, my jako istoty ludzkie sami jesteśmy częścią wszechświata i pozostajemy pod wpływem strzałki czasu, a więc dla nas kierunek upływu czasu także by się odwrócił. Po drugie, osiągnięcie klasycznego maksimum w ramach kosmosu kwantowego nie jest możliwe! Z logicznej analizy równania Wheelera-DeWitta i jego rozwiązań wynika bowiem, że wszechświat w obliczu zbliżania się do maksimum przechodzi w fazę kwantową, w której zarówno czas, jak i klasyczne przybliżenie znikają. W pewnym sensie dochodzi do interferowania ze sobą wielu odgałęzień, czyli wielu klasycznych światów częściowych funkcji falowej, czego efektem jest koniec klasycznej ewolucji. Ostatecznie powstaje ta sama bezczasowa mgła, którą widzimy również obecnie, spoglądając na wczesny rozwój wszechświata. W tym wypadku mamy do czynienia z sytuacją, w której zjawiska kwantowe są istotne także dla dużego wszechświata, a więc z dala od skali Plancka – kwantowy kosmos jest bardziej fundamentalny niż klasyczny wszechświat.

Zjawiska kwantowe w punkcie zwrotnym zapobiegają też powstawaniu paradoksów. W przypadku klasycznego odwrócenia strzałki czasu należałoby bowiem oczekiwać, że mogłaby do nas dotrzeć informacja z przyszłości – z okresu, w którym wszechświat się zapada, a czas płynie w odwrotnym kierunku. Najbardziej spodziewane byłoby jednak promieniowanie elektromagnetyczne, wyemitowane na przykład przez

gwiazdy pochodzące z przyszłości, a tym samym konfrontacja z nową odmianą paradoksu Olbersa (zob. część 4). „Mgła" powstała w wyniku działania zjawisk kwantowych w punkcie zwrotnym, a tym samym zanik czasu w tamtym punkcie sprawia, że taki scenariusz staje się niemożliwy. Kosmos kwantowy, zastępujący klasycznie odwrócony wszechświat, jest spójny. Posiada on wiele odgałęzień z wieloma obserwatorami, również niezliczoną ilość kopii nas samych, ograniczonych w swoim postrzeganiu wyłącznie do rozrastającego się świata, który w momencie zbliżania się do punktu zwrotnego kończy się przejściem do bezczasowości.

Na pytanie, czy taki scenariusz rzeczywiście opisuje naszą naturę, nie potrafimy jeszcze udzielić żadnej odpowiedzi. Brzmi on jednak logicznie, gdy połączymy ze sobą dwie sprawdzone teorie – teorię kwantową i ogólną teorię względności – oraz bez żadnych uprzedzeń prześledzimy konsekwencje wynikające z tego połączenia dla klasycznie odwróconego wszechświata.

Także w poważnej literaturze rozważa się niekiedy możliwość podróży w czasie. Ogólna teoria względności pozwala w zasadzie na istnienie rozwiązań umożliwiających podróż w czasie. Zazwyczaj są to bardziej złożone czasoprzestrzenie, jak na przykład „tunel czasoprzestrzenny", będący czymś w rodzaju skrótu między punktami w przestrzeni w formie kanału. Zgodnie ze szczegółowymi badaniami odpowiednie wykorzystanie takiego tunelu czasoprzestrzennego mogłoby przeistoczyć go w wehikuł czasu.

W przypadku podróży w czasie mogłoby dojść do powstania paradoksów przytaczanych w literaturze science fiction, na przykład: zabicie swojego dziadka w przeszłości uniemożliwiałoby własną egzystencję, a tym samym możliwość podjęcia podróży w czasie i w konsekwencji także zabicia dziadka. W ramach teorii względności możliwe są wyłącznie takie podróże w czasie, które wykluczają występowanie tego typu paradoksów – po prostu nie ma możliwości sformułowania spójnych warunków brzegowych, które prowadziłyby do powstania takich paradoksów. Włączenie entropii do scenariusza kosmosu

kwantowego powoduje dalsze umocnienie tego efektu. Wszyscy obserwatorzy są podporządkowani strzałce czasu i z tego powodu nigdy, również w przypadku istnienia tunelu czasoprzestrzennego, nie mogą cofnąć się do stanu o niższej entropii, czyli do przeszłości. W związku z tym musimy pogodzić się z faktem, że nigdy nie odwiedzi nas gość z przyszłości.

Zasada antropiczna

Czy warunkiem zrozumienia naszego świata jest uwzględnienie faktu naszej egzystencji? Kwestia ta jest w ostatnim czasie dość głośna, ponieważ nie wiadomo do końca, czy fundamentalne teorie definiują wszystkie stałe w przyrodzie, a zatem czy istnieje jeszcze miejsce na budowę świata. Pojęcie *zasady antropicznej* zostało wprowadzone w 1974 roku przez Brandona Cartera, fizyka teoretycznego pracującego w Meudon koło Paryża i specjalizującego się w teorii względności. Co kryje się za tym pojęciem? Zasada antropiczna ma być pewnego rodzaju kompromisem między dwiema innymi zasadami. Jedną z nich była zasada autocentrystyczna, przypisująca nam uprzywilejowane miejsce we wszechświecie. Wyobrażenie to najlepiej ucieleśniał przedkopernikański obraz świata z Ziemią umieszczoną w centrum wszechświata. Drugą z nich była natomiast silna interpretacja zasady kosmologicznej, omówionej przez nas w części 4 niniejszej książki, według której wszechświat wygląda wszędzie tak samo, a nasze otoczenie jest typowe dla każdego miejsca we wszechświecie. Carter słusznie zauważył, że silna zasada kosmologiczna nie może być uniwersalna, bo żyjemy przecież na wyjątkowej planecie pośrodku zimnego kosmosu. Carter postulował zatem, aby takie samo prawdopodobieństwo przypisywać nie miejscu, lecz wszystkim obserwatorom, którzy w dostatecznym stopniu są do nas podobni. Oczywiście nie wyklucza to również pozaziemskich form życia, a w związku z tym pojęcie zasady „antropicznej” nie do końca oddaje ideę, która za nim stoi, jednak taka nazwa się przyjęła i jest powszechnie używana.

Carter wprowadził dość ścisły podział między słabą a mocną wersją zasady antropicznej. Słaba zasada antropiczna dotyczy wyłącznie naszego wszechświata bądź jego modelu. Mocna zasada mówi z kolei o zespole możliwych wszechświatów, które różnią się od siebie na przykład pod względem wielkości stałych fizycznych.

Słaba zasada antropiczna mówi tylko lakonicznie, że parametry naszego świata muszą przyjmować wartości zgodne z naszą egzystencją. Brzmi to jak oczywistość i tak też jest, ale stwierdzenie to ma również ciekawe następstwa. Ważnym przykładem bezpośredniego wykorzystania słabej zasady antropicznej jeszcze przed jej sformułowaniem była krytyka wyrażona w 1961 roku przez amerykańskiego fizyka Roberta Dicke'a (1916–1997). Dotyczyła ona prezentowanego przez Paula Diraca wyobrażenia o zmieniającej się w czasie stałej grawitacyjnej. Dirac dziwił się, że w porównaniu z innymi oddziaływaniami, na przykład elektromagnetyzmem, grawitacja jest tak niewielka, i w 1937 roku przedstawił tezę, zgodnie z którą stała grawitacyjna maleje w czasie, a zatem nie jest stała. Dirac twierdził ponadto, że siła ciężkości jest tak niewielka, ponieważ wszechświat jest stosunkowo stary. Dicke przekonywał jednak, że zaobserwowana niewielka wartość stałej grawitacyjnej świadczy jedynie o istnieniu możliwości powstawania gwiazd i planet nadających się do zamieszkania także przy obecnym wieku świata. Wyłącznie w przypadku faktycznej stałej grawitacyjnej można bowiem powiedzieć, że obecny wiek wszechświata odpowiada mniej więcej długości życia gwiazdy ciągu głównego, na przykład Słońca. Współczesne badania zdecydowanie obaliły prognozę Diraca, jakoby stała grawitacyjna miała zmieniać się w czasie – wszelkie obserwacje przemawiają za tym, że jej wartość jest stała.

W nawiązaniu do stałych fizycznych stwierdzono z niemałym zdumieniem, że rozwój życia, takiego jakie znamy, jest możliwy tylko wtedy, gdy stałe te mają odpowiednio dobrane wartości. Silna zasada antropiczna wybiera spośród wszystkich możliwych wszechświatów z wszelkimi możliwymi wartościami tych stałych dokładnie te, które

zapewniają warunki sprzyjające rozwijaniu się życia. Często zdarza się, że zasada ta jest mylnie rozumiana jako podejście teleologiczne, według którego wszechświat od początku został tak zaprojektowany, aby mógł wydać na świat nas. Nie to mamy tu na myśli.

Również w przypadku silnej zasady antropicznej dysponujemy wczesnym przykładem jej zastosowania. Brytyjski astronom Fred Hoyle wraz ze swoimi współpracownikami stwierdził w 1953 roku, że pewien proces (tak zwany potrójny proces alfa), w ramach którego w gwiazdach zbudowanych z wodoru i helu powstają pierwiastki cięższe, zachodzi wyłącznie przy bardzo specyficznych wartościach odpowiednich stałych sprzężenia. Niestety w większości możliwych światów występują głównie wodór i hel, a zatem nie ma tam warunków niezbędnych do powstania życia – mało interesująca perspektywa.

Prawdopodobnie najbardziej znanym przykładem zastosowania silnej zasady antropicznej jest stała kosmologiczna. Zgodnie z tym, co mówiliśmy w części 4, stała ta przyjmuje w naszym wszechświecie niewielką wartość dodatnią, która jest jednak dużo niższa, niż przewidują wszystkie dotychczasowe prognozy. Zasada antropiczna zdaje się wyjaśniać, dlaczego tak się dzieje: zbyt wysoka wartość tej stałej spowodowałaby zbyt szybką ekspansję wszechświata, a co za tym idzie – niemożliwe byłoby powstanie galaktyk oraz życia. Gdyby przyjęła ona natomiast zbyt wysoką wartość ujemną, nie mogłyby powstać galaktyki, ponieważ wcześniej doszłoby do ponownego zapadania się wszechświata. Według obliczeń gęstość energii „ciemnej energii" (zob. część 4), która jest powiązana ze stałą kosmologiczną, pod żadnym pozorem nie może przekroczyć wartości sto razy większej od gęstości energii materii.

Wartość ta przekracza jednak znacznie wartość zaobserwowaną, która jest około trzy razy większa od gęstości energii materii. Czy istnieją zatem dodatkowe czynniki nieantropiczne? Kosmolog Aleksander Wilenkin zaostrzył kryteria zasady antropicznej, wprowadzając własną „zasadę przeciętności". Według niej znajdujemy się mniej więcej w środku obszaru wyznaczonego przez zasadę antropiczną, czyli jesteśmy dość typowi, po

prostu przeciętni. Na podstawie tego stwierdzenia za pomocą zwykłego szacunku Wilenkin zaproponował wartość gęstości energii stałej kosmologicznej, która oscyluje wokół dziesięciokrotności gęstości energii materii, a więc nie różni się znacznie od wartości wynikającej z obserwacji.

Carter stwierdził, że większość zastosowań zasady antropicznej prowadzi jedynie do potwierdzenia znanych już wyników. Są jednak wyjątki. Carter doszedł do wniosku, że istnienie istot inteligentnych we wszechświecie nawet na planetach o warunkach sprzyjających rozwojowi życia należy do rzadkości. Dlaczego? Weźmy za przykład nasze istnienie – nie trwa ono dużo krócej niż długość życia gwiazdy ciągu głównego, a zatem Słońce ma za sobą już ponad połowę swojej egzystencji. Jeśli uwzględnimy również, że nasze istnienie jest związane z wieloma znaczącymi przypadkami, rzeczywiście można dojść do wniosku o niewielkim prawdopodobieństwie istnienia życia. Z tego powodu Carter nie inwestowałby wielkich pieniędzy w programy badawcze mające odkryć życie pozaziemskie.

W przeszłości przemyślenia na temat zasady antropicznej nie były w fizyce zbyt popularne. Antropiczne wyjaśnienie danego problemu interpretowano zazwyczaj jako wyraz brakującej jeszcze prawdziwej wiedzy. Na początku XXI wieku karta się jednak odwróciła. Przełom nastąpił pod wpływem teorii strun, którą omówiliśmy w części 5, co może szczególnie dziwić, ponieważ właśnie przez nią rozważania antropiczne były wcześniej potępiane.

Pierwotnie teoria strun hołdowała zasadzie, że teoria jednoznacznie definiuje wszystkie parametry w fizyce. Według niej obliczenie niektórych parametrów nie jest wprawdzie jeszcze możliwe, ale lepsze zrozumienie teorii w przyszłości będzie potrafiło to zmienić. Ewentualność ta jest wciąż ważna, jednak wielu teoretyków zajmujących się teorią strun zdaje się zauważać, że teoria pozostawia bardzo wiele otwartych możliwości, na przykład umożliwia istnienie miriadów tak zwanych stanów podstawowych („próżni”), z których każdy opisuje możliwy świat. Obliczenia dopuszczają

istnienie nawet ponad 10^{500} takich próżni. Jak niewyobrażalnie ogromna to liczba, staje się jasne, gdy zdamy sobie sprawę, że łączna liczba wszystkich protonów i neutronów w widocznej dla nas części wszechświata wynosi „zaledwie” 10^{80}. Amerykański fizyk Leonard Susskind na określenie tej mnogości próżni wprowadził pojęcie *krajobrazu*. Można przez to rozumieć rozkład energii w przestrzeni pełnej wszelkich możliwych parametrów, w ramach którego występują możliwe stany podstawowe, takie jak niewielkie dolinki w pagórkowatym krajobrazie. Pojęcie to stosuje się również w dziedzinie biochemii, gdzie opisuje ono na przykład możliwości podczas zwijania białka.

Czy jednak wszystkie doliny w krajobrazie energetycznym są równoznaczne z możliwymi wszechświatami? Byłoby to bardzo zaskakujące, ponieważ oznaczałoby, że z ponad 10^{500} możliwości w cudowny sposób wybrana została akurat ta sprzyjająca powstaniu życia. W związku z tym uznaje się, że w rzeczywistości wszystkie 10^{500} możliwości jest realizowanych w formie miriadów wszechświatów będących częścią gigantycznego superświata. Pojęciem, które rozpowszechniło się w ostatnim czasie na określenie tego superświata, jest *multiświat*. Trzeba przyznać, że nazwa ta jest dość dziwna, gdyż już sam wszechświat obejmuje wszystko, nawet wtedy, gdy jest niewyobrażalnie duży i skomplikowany. Poszczególne wszechświaty częściowe mają odrębne przestrzenie lub czasoprzestrzenie oraz różnią się od siebie wartościami stałych fizycznych bądź nawet formą określonych praw natury. Każdy wszechświat częściowy może być przynajmniej tak duży jak widoczna dla nas część „naszego” wszechświata.

Nie trzeba koniecznie zajmować się teorią strun, aby spotkać się z terminem multiświata. Mianem multiświata określiliśmy na przykład w części 4 tej książki pewien model inflacyjnego wszechświata. W modelu tym powstają niezależne obszary częściowe o zróżnicowanych wartościach parametrów, tworząc wspólnie jeden multiświat. Nasz wszechświat powstaje na bazie takiego niewielkiego, w znacznej mierze jednorodnego obszaru częściowego (porównaj z rys. 4.13).

Z uwagi na to, że w przypadku multiświata chodzi o jeden prawdziwy świat, w obrazie tym silna zasada antropiczna zostaje zredukowana do słabej zasady antropicznej. Nie musimy się już zatem dłużej dziwić, że realizowany jest właśnie ten wszechświat, który umożliwia istnienie ludzi. Wszystkie możliwości dopuszczone przez tę teorię są urzeczywistnione, jednak pozostałych światów nigdy nie będziemy mogli zobaczyć, a my znajdujemy się po prostu w jednym z tych niewielu światów częściowych, w których panujące warunki sprzyjają rozwojowi życia w znanej nam formie.

Istnienie 10^{500} stanów podstawowych, z których każdy tworzy własny świat, wcale nie musi oznaczać braku wiarygodności teorii strun. Steven Weinberg uważa, że zasada antropiczna jest potrzebna wyłącznie do zdefiniowania kilku podstawowych parametrów, do których zalicza się stałą kosmologiczną i masę cząstki Higgsa. Na ich podstawie według Weinberga można obliczyć pozostałe wielkości. Gdyby założenie to okazało się prawdziwe, oznaczałoby to zaproponowanie przez fundamentalną teorię zupełnie nowej mieszanki stałych w fizyce, składającej się z czegoś policzalnego oraz czegoś nieokreślonego, ale opartego na założeniach antropicznych.

Dotychczasowe rozważania dotyczące zasady antropicznej w dużej mierze pomijają teorię kwantową, a uwzględnienie jej sprawi, że pojawią się dodatkowe struktury, które omawialiśmy w poprzednich rozdziałach tej książki. Do klasycznych stanów podstawowych dołączą ich możliwe superpozycje w kosmicznej funkcji falowej. W kontekście tym z uwagi na dekoherencję trzeba będzie również liczyć się z wystąpieniem wielu odgałęzień w klasycznej postaci. Dopiero uwzględnienie wszystkich odgałęzień ukazałoby pełen obraz kwantowego kosmosu.

Obraz świata antycznego był ujmująco prosty – Ziemia i wraz z nią człowiek w centrum nieskomplikowanego i bardzo harmonijnego kosmosu. Obecnie mamy do czynienia ze światem, którego wielkość i stopień skomplikowania są zatrważające. Nowy obraz świata nie jest jednak żadnym urojeniem, wręcz przeciwnie – powstaje on w wyniku

współdziałania dwóch dobrze sprawdzonych empirycznie teorii: kwantowej i względności. Obraz ten wynika zatem ze zgromadzonych doświadczeń. Proste i jednorodne są już tylko podstawowe pojęcia i zasady.

7. Refleksje

Jedynym problemem jest problem rzeczywistości; jest on w równym stopniu niemożliwy do rozwiązania, co aktualny.

Fernando Pessoa, *Księga niepokoju*

W styczniu 1957 roku na Uniwersytecie Karoliny Północnej zebrali się fizycy, aby podyskutować na temat rozwoju w dziedzinie teorii względności. Obecne były między innymi takie sławy jak Richard Feynman, John Wheeler oraz Bryce DeWitt. Feynman uczestniczył w dyskusji na palący zarówno wtedy, jak i teraz temat kwantyzacji grawitacji, wypowiadając się z charakterystyczną dla siebie szczerością. Podkreślał, że możemy wpaść w tarapaty, wierząc w mechanikę kwantową, a jednocześnie nie kwantyzując teorii grawitacji. Ponadto dodał, że jeżeli wierzy się w uniwersalność mechaniki kwantowej, to powinno się również wierzyć w grawitację kwantową.

Uniwersalność teorii kwantowej wiąże się rzecz jasna z aspektem empirycznym i właśnie z tego powodu eksperymenty dotyczące podstaw teorii kwantowej są tak istotne. Jak dotąd obowiązywanie jej zostało potwierdzone w każdym eksperymencie, na przykład w wielu doświadczeniach dotyczących stanów splątanych. Dodatkowo z formalizmu teorii kwantowej wynika czas powstania klasycznych właściwości oraz ich skala. Właśnie z tego powodu podczas dyskusji o komosie kwantowym wychodziłem z założenia o uniwersalności teorii kwantowej. Moim celem było ukazanie wpływu zastosowania tej teorii na wszechświat, nawet jeśli wpływu tego nie można *obecnie* bezpośrednio sprawdzić. Niezmiernie wrażliwe oddziaływanie na siebie układów kwantowych prowadzi nas nieuchronnie do pojęcia funkcji falowej wszechświata. Z uwagi na to, że wszechświat w dużych skalach jest

zdominowany przez grawitację, niezbędna w tym kontekście wydaje się teoria grawitacji kwantowej. Obecnie znane są jedynie jej założenia, jednak sięgają one tak daleko, że już teraz nietrudno nie zauważyć ogromnego wpływu, jaki wywrze taka teoria na współczesną fizykę.

Funkcja falowa wszechświata składa się z wielu odgałęzień, które z uwagi na dekoherencję posiadają właściwości niezależnych światów klasycznych. Zaledwie w niewielu tych odgałęzieniach panują warunki odpowiednie do tego, aby powstały w nich galaktyki, gwiazdy oraz życie w znanej nam formie. Wybór ten można opisać za pomocą słabej zasady antropicznej. Rzuca to zupełnie nowe światło na stwierdzenia Leibniza o tym, że żyjemy w najlepszym ze wszystkich możliwych światów.

W tym miejscu przychodzą oczywiście na myśl pytania o ogólne znaczenie i interpretację tych wyobrażeń. Już w przypadku zwykłej mechaniki kwantowej istnieją przecież różne podejścia do kwestii interpretacji. W pewnym uproszczeniu można więc mówić o dwóch nurtach: *pozytywizmie* oraz *realizmie*. (Pominiemy tu skrajne stanowisko solipsyzmu, według którego istnieje tylko jeden obserwator).

Są oczywiście różne odmiany pozytywizmu. Pierwotnie oznaczał on po prostu ograniczenie do twierdzeń nauk przyrodniczych. Znaczenie pozytywizmu w fizyce i jego związek z realnym światem zewnętrznym krótko podsumował Max Planck w swoim wykładzie z 1930 roku. Według niego zadanie fizyki można scharakteryzować w sposób wyczerpujący, mówiąc, że polega ono na tym, aby do opisu różnorodnych istniejących obserwacji przyrody wprowadzić możliwie dokładny i nieskomplikowany kontekst prawny. Na każde pytanie, które pozytywizm w ogóle uznaje za uzasadnione, można zatem udzielić odpowiedzi na podstawie obserwacji. Prawdopodobnie najlepszym odzwierciedleniem koncepcji pozytywizmu była ekonomia myślenia skonstruowana przez Ernsta Macha, według której opis przyrody sprowadzał się do przedstawienia jej zjawisk w odniesieniu do ekonomii. Koncepcja ta odrzucała istnienie realnego świata rzeczywistego. Podczas swojego wykładu Planck wyraził opinię, że pozytywista dostrzegłby w Koperniku nie rewolucyjnego odkrywcę, lecz

genialnego wynalazcę, ponieważ dostarczył jedynie bardziej ekonomicznego opisu tych samych obserwacji astronomicznych.

Sam Planck stał w aktywnej opozycji do tego nurtu. W tej kwestii był on zgodny z innymi fizykami, takimi jak Einstein, Schrödinger czy Pauli, mówiąc, że największym niebezpieczeństwem, jakie niesie ze sobą pozytywizm, jest jego jałowość. W swoim stylu do tematu odniósł się Einstein, pisząc w liście do Besso z 13 maja 1917 roku, że co prawda nie przeklina rumaka Macha, ale nie uważa też, aby sprowadził on na świat coś nowego poza wytępieniem szkodliwego robactwa. Nic więc dziwnego, że fizycy o nastawieniu pozytywistycznym raczej nie zajmują się grawitacją kwantową czy kosmologią kwantową. Pozytywiści nie interesują się tymi problemami, dopóki nie istnieją obserwacje, które rzeczywiście potrzebowałyby ich opinii. Oczywiście jednocześnie takim nastawieniem sami pozbawiają się szansy odkrycia czegoś nowego.

Zgodnie z opinią Schrödingera pozytywizm jest z pewnością nurtem o niepodważalnych fundamentach logicznych, jednak jego samodzielnie wytyczone granice ograniczają go w dążeniu do odkrycia czegoś nowego w miejscach, w których istnieje taki potencjał. Znanym przykładem tego może być odrzucenie przez pozytywistów istnienia atomu na początku XX wieku.

W naturalny sposób przechodzimy do przeciwwagi pozytywizmu – realizmu. W swoim wykładzie Planck określał go mianem hipotezy mówiącej o tym, że nasze przeżycia nie są istotą świata fizycznego, a raczej informują nas o istnieniu innego świata, który za nimi stoi i jest od nas niezależny, czyli innymi słowy, że istnieje realny świat zewnętrzny. W żadnym razie nie należy przez to rozumieć naiwnego realizmu oznaczającego to, że od razu rozpoznajemy istniejący naprawdę świat zewnętrzny jako rzecz samą w sobie. Wprawdzie zakładamy obecność niezależnej rzeczywistości, jednak wychodzimy z założenia, że jej istnienia nie można udowodnić, a zatem nigdy nie możemy mieć co do niej absolutnej pewności. Fizyk kwantowy Bernard d'Espagnat mówi na

przykład o ukrytej rzeczywistości, do której można się tylko zbliżyć, ale nigdy ją osiągnąć.

Oczywiście wypowiedzi realistów również muszą wyrażać harmonię między eksperymentami a obserwacjami. W tym względzie wyznają oni tę samą opinię co pozytywiści. Poza samymi wzorami empirycznymi pozytywiści mają jednak pełną swobodę w kwestii ustalania norm i snucia domysłów. Einstein podkreślał, że połączenie między postrzeżeniem zmysłowym a opisem matematycznym nie ma charakteru logicznego, lecz opiera się na intuicji: „Pojęcia fizyczne są swobodnymi tworami umysłu ludzkiego i nie są, choć by się tak mogło wydawać, w sposób jednoznaczny wyznaczone przez świat zewnętrzny" (A. Einstein oraz L. Infeld, *Ewolucja fizyki*, wg przekładu Ryszarda Gajewskiego, 1998). Zgodnie z opinią Einsteina wyrażoną w *Autobiographisches* dopiero stopień pewności, z jaką można utworzyć ten związek, bądź intuicyjne powiązanie odróżniają pustą fantastykę od naukowej „prawdy". Chciałbym, aby moja książka była rozumiana w ten sposób.

Z perspektywy realizmu prawa natury się odkrywa, a nie wynajduje. W pierwszej chwili zdumiewa przy tym ich prostota. Właściwie dlaczego można opisać te prawa za pomocą względnie nieskomplikowanych struktur matematycznych? Eugene Wigner w nawiązaniu do tego tematu mówił o nierozsądnej przydatności matematyki do opisu przyrody. Ostatecznej odpowiedzi prawdopodobnie nigdy nie otrzymamy, lecz na pewno możemy posługiwać się zasadą antropiczną. Istnienie życia z całą pewnością zakłada obecność regularności (struktur), na przykład w formie symetrii. Matematyka to jednak nauka, która jest stworzona do tego, aby opisywać ogólne struktury. W rzeczywistości współczesna fizyka opiera się na symetriach, które można opisać za pomocą prostych reguł matematycznych.

W dziedzinie matematyki zdania co do tego, czy wzory matematyczne się odkrywa, czy wynajduje, są również podzielone. W odróżnieniu od praw natury moim zdaniem struktury matematyczne się wynajduje. Z uwagi na to, że matematykę często porównuje się do muzyki (równie

często zdolności te występują paralelnie), można także powiedzieć, że prawa matematyczne się „komponuje". Hermann Weyl jest zdania, że być może „matematyzowanie", podobnie jak muzykowanie, stanowi czynność twórczą człowieka, której efekty są uzależnione nie tylko pod względem formy, ale też treści, od decyzji podejmowanych w przeszłości i dlatego nie można ich pojąć w sposób całkowicie obiektywny. Szwajcarski pisarz Friedrich Dürrenmatt proponuje dodatkowo, aby „być może" zastąpić „przede wszystkim". Ma rację.

Oczywiście można wierzyć, że struktury matematyczne znajdują się w gigantycznej bibliotece w Babilonie i tylko czekają na odkrycie. Prawdopodobieństwo, że rzeczywiście tak jest, jest jednak tak samo niskie jak w przypadku książek ze słynnego opowiadania Jorge Luisa Borgesa. W nawiązaniu do muzyki można by zatem, zgodnie z powyższym porównaniem, wyobrazić sobie, że na przykład symfonia „Jowiszowa" Mozarta istniała już w takim układzie nut i tylko czekała na swojego odkrywcę. Mało kto odważyłby się powiedzieć, że ta wizja nie jest absurdalna. Dlaczego więc w przypadku matematyki miałoby być inaczej?

Być może twórcza działalność w dziedzinie matematyki nie służy przede wszystkim udowodnieniu twierdzeń, lecz znalezieniu odpowiednich struktur i definicji potrzebnych do skonstruowania ciekawych praw. Bardzo często bodźca potrzebnego do powstania takich struktur dostarcza właśnie fizyka. Newton odkrył rachunek różniczkowo-całkowy głównie po to, aby opisać ruch ciał rzeczywistych.

Filozof Ludwig Wittgenstein doszedł do wniosku, że dyskusja na tematy fundamentalne w zakresie matematyki jest w gruncie rzeczy zbyteczna i podejmuje się ją tylko dlatego, że zapomniano, do czego ona tak naprawdę służy. W jednym z interesujących nas fragmentów pisze on na przykład:

> Nieporozumienia w tych sprawach [dyskusji na tematy fundamentalne – przyp. autora] wynikają z tego, że matematykę traktuje się jako rodzaj nauki przyrodniczej. A to z kolei wiąże się z tym, że matematyka odłączyła się od nauk przyrodniczych, ponieważ dopóki uprawia się ją w bezpośrednim połączeniu

z fizyką, pozostaje jasne, że nie jest *ona* żadną nauką przyrodniczą. (Podobnie jak w przypadku miotły, której nie uznaje się za mebel, dopóki służy do zamiatania, a nie jest jednym z elementów wystroju pomieszczenia).

Zjawisko to jest nam znane ze współczesnej fizyki, a w szczególności z badań nad grawitacją kwantową, gdzie istnieją nurty, które traktują miotłę jak element wystroju pomieszczenia i inwestują wiele wysiłku w dalszy rozwój języka matematycznego zamiast zadbać o odniesienie do świata zewnętrznego. W przypadku, gdy brakuje intuicyjnego powiązania ze światem, co postulował Einstein, grozi nam niebezpieczeństwo pogrążenia się w przywołanej przez niego pustej fantastyce.

Wielokrotnie nawiązywaliśmy do kwestii położenia człowieka jako obserwatora w kosmosie kwantowym. Czy przedstawiony tu scenariusz może rzucić nowe światło na odwieczny problem istnienia wolnej woli? Przypomnijmy sobie najpierw, o co chodziło: jeśli założymy istnienie determinizmu praw natury, to okaże się, że wszystko, co się dzieje, łącznie z naszymi działaniami, zależy od warunków początkowych. Oznaczałoby to, że wolna wola nie istnieje, a to z kolei wiązałoby się z nieprzewidywalnymi konsekwencjami dla moralności, prawa karnego i praktycznie wszystkich innych aspektów ludzkiego współżycia. Z drugiej strony uczucie posiadania wolnej woli jest przecież decydujące dla naszej egzystencji.

Jak przedstawia się ten dylemat z perspektywy kosmosu kwantowego? Przecież już podczas dyskusji na temat teorii kwantowej wyodrębniliśmy perspektywę ptaka i żaby. Perspektywa ptaka odpowiada hipotetycznemu spojrzeniu z zewnątrz na świat kwantowy, które ukazuje go jako świat bezczasowy, pozbawiony wszelkich klasycznych właściwości, a zatem w tym przypadku pytanie o istnienie wolnej woli w ogóle się nie pojawia. Zupełnie inaczej wygląda wewnętrzna perspektywa żaby, w której istnieje niezliczona ilość klasycznych komponentów wraz z obserwatorami i rozwojem w czasie. Z uwagi na to, że nie możemy z góry założyć uniwersalnej tożsamości obserwatora, rozważanie determinizmu jego działań jest także pozbawione sensu. Z perspektywy naszego niewielkiego

zakamarka świata kwantowego możemy stale podtrzymywać wyobrażenie o istnieniu wolnej woli, ponieważ wolna wola i struktura kosmosu kwantowego są ze sobą kompatybilne. Co możemy powiedzieć na temat świadomości? Jej opis musi oczywiście współgrać ze wszystkimi prawami natury. Bezpośrednie doświadczenie pozostanie jednak na zawsze subiektywne – wystarczy tylko przypomnieć sobie znany przykład postrzegania barw. Również pojęcie teraźniejszości jest pojęciem subiektywnym, które nie znajduje żadnego odpowiednika w prawach natury.

Nasza hipotetyczna podróż pociągiem z Paryża do Lizbony powoli dobiega końca. Zrobiło się późno, a nasza rozmowa zbacza w kierunku ogólnych obserwacji. Na koniec pozwólmy, aby jeszcze raz do głosu doszedł profesor Katschka: „Przemijalność mianowicie nie tylko nie pozbawia wartości, lecz wprost przeciwnie, to ona właśnie nadaje wszelkiemu istnieniu wartość, godność i wdzięk. Tylko to, co epizodyczne, co ma swój zaczątek i kres, jest zajmujące i wzbudza sympatię, będąc przepojone nietrwałością (...). Dobrej nocy!”.

Uwagi

Do części 1

Od starego kosmosu do przestrzeni i czasu

Cytat Hansa Blumenberga pochodzi z jego wspaniałego wstępu do dzieła Galileusza *Sidereus Nuncius* (*Nachricht von neuen Sternen*) (Suhrkamp, Frankfurt am Main 1980). Uwaga Einsteina na temat Newtona pochodzi z artykułu: A. Einstein, *Newtons Mechanik und ihr Einfluß auf die Gestaltung der Theoretischen Physik*, przedrukowanego w *Mein Weltbild* (Ullstein, Frankfurt am Main 1977).

Od przestrzeni i czasu do czasoprzestrzeni: szczególna teoria względności Einsteina

Fragment dotyczący eksperymentu myślowego, opracowanego przez Einsteina w szkole kantonalnej w Aarau, pochodzi z książki: *Helle Zeit – dunkle Zeit* pod redakcją C. Seeliga (Europa-Verlag, Zürich 1956), s. 10. Wiele ciekawych informacji na temat czasów szkolnych Einsteina w Aarau można znaleźć w książce *Der jugendliche Einstein und Aarau* pod redakcją H. Hunzikera (Birkhäuser, Basel 2005).

Opis wspomnianego w tekście, jak dotąd najdokładniejszego testu wzoru Einsteina $E = mc^2$ znajduje się w: S. Rainville i in., *World Year of Physics: A direct test of* $E = mc^2$, „Nature" 438, 1096–1097 (2005).

W odniesieniu do czasoprzestrzeni Minkowskiego: Minkowski bardzo energicznie rozpoczął swój wykład w Kolonii, mówiąc: „Moi Panowie! Spostrzeżenia o przestrzeni i czasie, które pragnę Wam przedstawić, opierają się na eksperymentach fizycznych i to jest ich mocna strona. Zmiana kierunku jest radykalna. Od tej chwili przestrzeń i czas jako

osobne zjawiska odejdą w cień, a tylko połączenie tych dwóch zachowa niezależność".

Geometryzacja świata: ogólna teoria względności Einsteina

W 1916 roku Einstein w taki sposób podsumował najważniejsze etapy na drodze do sformułowania ogólnej teorii względności: „1907 – pomysł na powstanie ogólnej teorii względności; 1912 – uznanie nieeuklidesowej natury metryki i jej fizycznego uwarunkowania od grawitacji; 1915 – równania pola grawitacyjnego. Wyjaśnienie ruchu peryhelium Merkurego".

W odniesieniu do zasady równoważności: planowana misja kosmiczna STEP (*Satellite Test of the Equivalence Principle*) powinna sprawdzić zasadę równoważności z dokładnością 10^{-18}. Źródłem ewentualnych niezgodności z tą zasadą mogą być teorie zunifikowane, jak na przykład teoria strun (zob. część 5). Obecnie nie ma jednak pewności co do tego, kiedy (i czy) satelita ten wystartuje. Lepsze rokowania ma francuski projekt kosmiczny MICROSCOPE, który ma się rozpocząć prawdopodobnie pod koniec 2011 roku. Jego zadaniem będzie zbadanie zasady równoważności z dokładnością około 10^{-15}. W przygotowaniach do tego projektu wykorzystywana jest między innymi wieża zrzutowa w Bremie.

Teoria względności w Układzie Słonecznym

Aktualny wykaz wszystkich testów teorii względności w obrębie Układu Słonecznego (i poza nim) można znaleźć w: D. Giulini, *Einsteins Kunstwerk*, „Physik-Journal" 4, 27–33 (2005) i cytowanej tam literaturze.

W odniesieniu do obrotu peryhelium Merkurego: europejska sonda do badania Merkurego BepiColombo (na 2013 r. planowany jest jej start, a na 2019 r. dotarcie na miejsce) pozwoli precyzyjnie określić położenie tej najbardziej wewnętrznie zlokalizowanej planety, a tym samym w setną

rocznicę sformułowania ogólnej teorii względności dostarczy nam jeszcze dokładniejszej wartości obrotu peryhelium.

Fale grawitacyjne

W tekście pojawiła się nazwa planowanej misji kosmicznej LISA, która według aktualnego harmonogramu powinna wystartować w 2018 roku. Poza misją LISA istnieją również wizje innych projektów, jak na przykład *Big Bang Explorer*. Satelita ten powinien rozpocząć swoją misję w 2025 roku, a jej celem będzie wykrycie fal grawitacyjnych powstałych w bardzo wczesnym okresie rozwoju wszechświata. Do powstania takich fal mogło dojść w fazie inflacyjnej (zob. część 4), czyli około 10^{-35} sekundy po wielkim wybuchu. Od tej pory rozprzestrzeniały się one w dużej mierze bez przeszkód we wszechświecie – prawdopodobnie żadne inne obserwacje nie mogłyby nas bardziej przybliżyć do wielkiego wybuchu!

Emisja fal grawitacyjnych została dowiedziona także pośrednio w odległym kwazarze: zob. M.J. Valtonen i in., „Nature" 452, 851–853 (2008). W kwazarze OJ287, odległym o 3,5 miliarda lat świetlnych, dookoła siebie krążą najprawdopodobniej czarna dziura o 18 miliardów razy większej masie od masy Słońca oraz czarna dziura o 100 milionów razy większej masie od masy Słońca. Większą dziurę otacza pierścień pyłowy, z którym w regularnych odstępach czasu zderza się mniejsza dziura, wywołując przy tym rozbłyski światła. Dzięki temu można określić ruch orbitalny. Da się zaobserwować skracanie okresu orbitalnego, co można wyjaśnić emisją fal grawitacyjnych. Autorzy doszli do wniosku, że do ostatniego rozbłysku światła doszłoby 20 dni później, gdyby nie było fal grawitacyjnych. Einstein byłby zachwycony!

Czarne dziury

Historia czarnych dziur zaczyna się na wiele lat przed teorią względności. Już pod koniec XVIII wieku pastor John Michell w Anglii, a potem Pierre-Simon Laplace we Francji przypuszczali, że mogą istnieć „ciemne

gwiazdy", w przypadku których ze względu na ich dużą masę druga prędkość kosmiczna przekracza prędkość światła. W tamtych czasach wyobrażenie to nie było niedorzeczne, ponieważ uważano, podążając za Newtonem, że światło składa się z niewielkich cząstek. Wraz z pojawieniem się falowej teorii światła na początku XIX wieku prace te zostały zapomniane, bo fale światła w mechanice Newtona nie podlegają działaniu grawitacji. Dopiero ogólna teoria względności ponownie poddała światło działaniu siły ciężkości, gdyż grawitacja jest tylko wyrazem geometrii przestrzeni i czasu, a przecież wszystko podlega grawitacji. Proponowane przez nią czarne dziury nie są jednak „czarnymi gwiazdami", lecz dziwną geometrią samej czasoprzestrzeni, w której część nie znajduje się w związku przyczynowym z pozostałością. Ciekawe informacje na temat czarnych dziur można znaleźć na przykład w książce Kipa Thorne'a (zob. spis literatury).

Ile wymiarów ma świat?

Bardzo trafną biografię Theodora Kaluzy i jego osiągnięć przedstawia: D. Wünsch, *Der Erfinder der 5. Dimension – Theodor Kaluza. Leben und Werk* (Termessos-Verlag, Göttingen 2007).

Do części 2

Co opisuje teoria kwantowa?

W odniesieniu do kongresu Solvaya w 1927 roku: zachował się nawet krótki film niemy przedstawiający głównych bohaterów tego kongresu. Film ten jest dostępny w Internecie pod adresem: http://www.maxborn.net/index.php?page=filmnews.

Stany splątane i obalenie rzeczywistości lokalnej

W odniesieniu do doświadczeń Younga: o doświadczeniu Younga z elektronami oraz dziewięciu innych z dziesiątki najpiękniejszych

eksperymentów wszech czasów można przeczytać w: *Die Top Ten der schönsten physikalischen Experimente* autorstwa C. Jönssona i A. Fässlera (Rowohlt, Reinbek 2005).

Wspomniany w tekście eksperyment Younga opisał F. Lindner i in. w *Attosecond double-slit experiment*, „Physical Review Letters" 95, artykuł nr 040401 (2005). Za szczelinę w eksperymencie służą jedno lub dwa okna, które mogą być otwarte na czas jednej attosekundy (10^{-18} sekundy) w celu fotojonizacji atomów. Dzięki temu można zaobserwować sytuację, w której elektron natrafia na pojedynczą i podwójną szczelinę, tworząc charakterystyczny wzorzec interferencyjny.

W odniesieniu do problemu EPR: twórcami problemu są A. Einstein, B. Podolsky oraz N. Rosen, którzy opisali go po raz pierwszy w publikacji *Can quantum-mechanical description of physical reality be considered complete?*, „Physical Review" 47, 777–780 (1935). Einstein stwierdził później, że uczoność tej pracy przesłoniła jej sens. Główne założenia pracy przedstawił on w sposób zrozumiały (na tyle zrozumiały, na ile można coś wyrazić w języku ojczystym) w artykule *Quanten-Mechanik und Wirklichkeit*, wydanym w „Dialectica" 2, 320–324 (1948). W artykule tym tak pisze o swoim założeniu lokalności: „Ponadto w przypadku tego uporządkowania rzeczy wprowadzonych do fizyki kluczowa wydaje się kwestia tego, że w określonym czasie rzeczy te zakładają swoją niezależną wzajemną egzystencję, o ile znajdują się one «w różnych częściach przestrzeni». Bez założenia takiej niezależności wzajemnej egzystencji («istoty») rzeczy odległych od siebie w przestrzeni, które wywodzi się z życia codziennego, nie byłoby możliwe myślenie fizyczne w znanej nam formie". Czynnik ten nie pasuje rzecz jasna do nielokalnego charakteru teorii kwantowej. To, że na poziomie bardziej fundamentalnym nie może istnieć żadna teoria lokalna, która odzwierciedlałaby wyniki teorii kwantowej, jest związane ze złamaniem nierówności Bella, ale tego Einstein nie mógł jeszcze wiedzieć.

W odniesieniu do nierówności Bella i rzeczywistości nielokalnej: oprócz artykułów naukowych John Bell napisał również kilka przystępniejszych

prac na ten temat. Zebrano je w publikacji: J.S. Bell, *Speakable and Unspeakable in Quantum Mechanics: Collected Papers on Quantum Philosophy*, wydanie drugie (Cambridge University Press, Cambridge 2004).

Nierówności Bella przedstawiają możliwe zależności między wynikami pomiarów w układach częściowych w przypadku założenia istnienia rzeczywistości lokalnej. Rzeczywistość lokalna oznacza, że układ posiada pewną właściwość (na przykład jedną składową spinu w dowolnym kierunku), a zdarzenia, do których dochodzi w innych odpowiednio oddalonych układach, nie mają na nią żadnego wpływu. W eksperymentach obserwuje się na przykład dwie rozchodzące się w powietrzu cząstki, na których skoncentrowana jest uwaga dwóch detektorów. Detektory mogą mierzyć spin w odniesieniu do dowolnego kierunku w przestrzeni. W celu zapewnienia lokalności odpowiednie kierunki muszą być wylosowane tak szybko, aby nie mogło dojść do wymiany sygnałów między nimi, a zatem różnica czasu między dwoma ustawieniami musi być mniejsza lub równa ilorazowi prędkości światła i różnicy odległości. Zaobserwowane doświadczalnie złamanie nierówności zaprzecza założeniu istnienia rzeczywistości lokalnej i odpowiada prognozom nielokalnej teorii kwantowej.

Wspomniana w książce praca Zeilingera i jego współpracowników to: S. Groeblacher i in., *An experimental test of non-local realism*, „Nature" 446, 871–875 (2007).

W odniesieniu do eksperymentu opóźnionego wyboru: John Wheeler opisuje ten pomysł w swoim artykule *Law without law*, zamieszczonym w książce *Quantum Theory and Measurement*, napisanej przez J.A. Wheelera oraz W.H. Zurka (Princeton University Press, Princeton 1983), s. 182–183. Pisze on: „Zapada zatem decyzja, czy foton pokonuje jedną, czy dwie drogi, gdy jest już w ruchu". Niedorzeczność tego zdania znika, gdy zrezygnujemy z wyobrażenia fotonu jako klasycznej kulki.

Opis wspomnianego w tej książce eksperymentu przeprowadzonego pod przewodnictwem Jeana-François Rocha można znaleźć w: V. Jacques i in.,

Experimental Realization of Wheeler's Delayed-Choice Gedanken Experiment, „Science" 415, 966–968 (2007).

Informacja kwantowa

W odniesieniu do rozkładu na czynniki pierwsze: jak dotąd nie powstał jeszcze żaden komputer kwantowy, który potrafiłby rozłożyć na czynniki pierwsze wysokie liczby stosowane podczas przesyłania zaszyfrowanych wiadomości. Co ciekawe, rozkłady trochę mniejszych liczb można przeprowadzić za pomocą pomysłowych eksperymentów fizycznych z wykorzystaniem interferencji. W pracy M. Gilowskiego i in., *Gauss Sum Factorization with Cold Atoms*, „Physical Review Letters" 100, artykuł nr 030201 (2008), dokonano na przykład eksperymentalnego rozkładu liczby 263 193 na czynniki pierwsze 3, 7, 83 oraz 151.

Jak powstaje klasyczny świat?

W odniesieniu do kota Schrödingera: oryginalny tekst napisany przez Schrödingera można znaleźć w przedstawionej poniżej pozycji, w której znajdują się również inne warte uwagi prace na temat teorii kwantowej: K. Baumann oraz R.U. Sexl, *Die Deutungen der Quantentheorie* (Vieweg, Braunschweig 1987).

W odniesieniu do dekoherencji: klasyczną pracą na temat lokalizacji obiektów makroskopowych poprzez oddziaływanie z otoczeniem jest: E. Joos oraz H.D. Zeh, *The emergence of classical properties through interaction with the environment*, „Zeitschrift für Physik" B 59, 223–243 (1985). Wspomniane w tekście obliczenia dotyczące dekoherencji dla procesów zachodzących w mózgu ukazały się w: M. Tegmark, *Importance of quantum decoherence in brain processes*, „Physical Review" E 61, 4194–4206 (2000); zob. również http://arxiv.org/abs/quant-ph/9907009.

Interpretacje teorii kwantowej

Więcej informacji na temat teorii Bohma można znaleźć na przykład w: D. Dürr, *Bohmsche Mechanik als Grundlage der Quantenmechanik* (Springer, Berlin 2001). W kwietniowym (2008) wydaniu czasopisma „Spektrum der Wissenschaft" zamieszczono artykuł na temat interpretacji Everetta, opisujący również skomplikowany życiorys Hugh Everetta.

Teoria kwantowa i cząstki elementarne

W odniesieniu do antycząstek: sformułowanie równania skłoniło Diraca do przedstawienia koncepcji antycząstek. W rzeczywistości jednak już Einstein kilka lat wcześniej udowodnił, że teoria, która połączy ogólną teorię względności i elektromagnetyzm, wymaga dla każdej naładowanej cząstki antycząstki o ładunku przeciwnym: zob. A. Einstein, *Elektron und Allgemeine Relativitätstheorie*, „Physica" 5, 330–334 (1925). Jego argument miał naturę klasyczną i wymagał tylko symetrii względem odwrócenia czasu. Z uwagi na to, że w 1925 roku wydawało się jeszcze, iż elektron o ładunku dodatnim nie istnieje, Einstein wątpił przez krótki czas w możliwość połączenia elektrodynamiki i teorii grawitacji.

W odniesieniu do teorii cechowania: w ramach teorii cechowania rozpatruje się przede wszystkim wielkości, które można przekształcić (poddać transformacji cechowania) niezależnie od siebie w każdym punkcie czasoprzestrzeni, nie zmieniając przy tym wielkości obserwowanych. Tę lokalną niezmienniczość cechowania wprowadził Hermann Weyl w 1929 roku. W tym samym czasie Einstein zajmował się alternatywnym ujęciem zwanym „teleparalelizmem", w ramach którego wielkości mimo dzielących je znacznych odległości są ze sobą nierozerwalnie powiązane. W swoim liście do Pascuala Jordana z 30 listopada 1929 roku Pauli nie odnosił się do tej idei entuzjastycznie: „Podobno Einstein na sympozjum w Berlinie opowiadał straszne bzdury o nowym teleparalelizmie! Sam fakt, że jego równania w ogóle nie przypominają teorii Maxwella, chce przyjąć za argument przemawiający za ich związkiem z teorią kwantową. Takimi bzdurami można zaimponować najwyżej amerykańskim dziennikarzom, a nie

amerykańskim fizykom, o europejskich nawet nie wspominając". Czasy się zmieniły!

W odniesieniu do mas cząstek: w fizyce cząstek elementarnych masy mierzy się najczęściej w elektronowoltach (a dokładniej: elektronowolt/c^2). Jeden elektronowolt (w skrócie: eV) odpowiada energii kinetycznej elektronu poddanego przyspieszeniu w potencjale elektrycznym o wartości jednego wolta (1 V) – 1000 eV to 1 keV (kiloelektronowolt), jeden milion eV to 1 MeV (megaelektronowolt), a jeden miliard eV to 1 GeV (gigaelektronowolt). W przeliczeniu na dżula (J) 1 eV to zaledwie $1{,}6 \times 10^{-19}$ J. Zgodnie ze wzorem $E = mc^2$ energie te odpowiadają również masie. W jednostkach tych masa elektronu wynosi około 511 keV, a masa protonu to około 1 GeV. Wspomniana w tekście cząstka Higgsa, którą mamy nadzieję znaleźć dzięki LHC, powinna mieć masę nie większą niż 150 GeV. Dzięki obserwacjom kosmologicznym (za pomocą satelity WMAP, o którym wspomnieliśmy w części 4) wiemy, że suma mas wszystkich rodzajów neutrin będzie wynosić poniżej 0,6 eV. Dokładne wartości mas w dalszym ciągu pozostają dla nas zagadką. Więcej na temat LHC można znaleźć w artykule *Der Large Hadron Collider – die Entdeckungsmaschine* autorstwa T. Müllera, wydanym w czasopiśmie „Physik in unserer Zeit", luty 2008, s. 78–85. Wiele ciekawych informacji dostarcza również pozycja: *Am Rande der Dimensionen. Gespräche über die Physik am CERN*, której autorem jest R. Landua (Suhrkamp, Frankfurt am Main 2008).

Promieniowanie czarnych dziur

Wspomnianą w tekście temperaturę czarnej dziury („temperatura Hawkinga") w przypadku sferycznie symetrycznej czarnej dziury (czyli czarnej dziury Schwarzschilda) o masie M można wyliczyć ze wzoru:

$$T = \frac{\hbar c^3}{8\pi k_{\mathrm{B}} G M} \approx 6{,}17 \times 10^{-8} \left(\frac{M_\odot}{M} \right) \text{ kelwin}$$

gdzie $M_\odot$ oznacza masę Słońca; ħ = h/2π, a k_B to stała Boltzmanna (porównaj z częścią 3). Jak widać, temperatura czarnych dziur o masie kilka razy większej od masy Słońca lub wyższej jest mikroskopijna. Temperaturę Hawkinga można zatem zmierzyć wyłącznie dla dziur o bardzo niewielkiej masie.

Do części 3

Mikroskopowa interpretacja entropii

Najważniejsze prace Boltzmanna na temat entropii zostały zebrane w książce: L. Boltzmann, *Entropie und Wahrscheinlichkeit* (Harri Deutsch, Frankfurt am Main 2000).

Entropia czarnych dziur

Słynny wzór Bekensteina i Hawkinga na entropię czarnych dziur to:

$$S_{\mathrm{BH}} = k_{\mathrm{B}} \frac{A}{4 l_{\mathrm{P}}^2} ,$$

gdzie k_B oznacza stałą Boltzmanna, l_P – długość Plancka (zob. część 5), natomiast A – powierzchnię horyzontu zdarzeń. W przypadku sferycznie symetrycznej czarnej dziury (czyli czarnej dziury Schwarzschilda) wzór na entropię wygląda następująco:

$$S_{\mathrm{BH}} \approx 1{,}07 \times 10^{77} k_{\mathrm{B}} \left(\frac{M}{M_\odot} \right)^2 ,$$

gdzie $M_\odot$ oznacza masę Słońca.

Do części 4

Zasada kosmologiczna

Aktualne i ugruntowane informacje podstawowe z zakresu naszego współczesnego stanu wiedzy na temat *cosmic web*, rozkładu świecącej i ciemnej materii we wszechświecie przekazuje czasopismo „Science" w wydaniu z 4 stycznia 2008 roku.

Cytat Carla Wirtza pochodzi z jego artykułu *De Sitters Kosmologie und die Radialbewegungen der Spiralnebel*, opublikowanego w czasopiśmie „Astronomische Nachrichten" 222, 21–26 (1924).

Topologia wszechświata

Wspomniany w tekście artykuł na temat dwunastościennego kształtu wszechświata to: J.-P. Luminet i in., *Dodecahedral space topology as an explanation of weak wide-angle temperature correlations in the cosmic microwave background*, „Nature" 425, 593–595 (2003). Warto przeczytać również komentarz George'a Ellisa na s. 566 tego numeru. Badanie N.J. Cornisha i in., *Constraining the Topology of the Universe*, „Physical Review Letters" 92, artykuł nr 201302 (2004), nie potwierdziło tych wyników. Autorzy uważają, że w przypadku znacznej klasy modeli możliwa skala powtórzenia musi przekraczać 24 Gpc.

Więcej informacji wprowadzających na temat właściwości przestrzeni trójwymiarowych można znaleźć na przykład w książce: D.B. O'Shea, *The Poincaré Conjecture: In Search of the Shape of the Universe* (Walker & Company, New York 2007). Książka ta powstała z okazji wspomnianego w tekście potwierdzenia hipotezy Poincarégo przez Grigorija Perelmana. Bardziej fachowych informacji na temat kształtu przestrzeni trójwymiarowych dostarczy książka J.R. Weeksa *The Shape of Space* (Marcel Dekker, New York–Basel 1985) oraz praca J.-P. Lumineta *Topologie et cosmologie*, zamieszczona w *Géometrie au XX^{e} siècle* pod redakcją J. Kouneihera i in., s. 351–361 (Hermann Éditeurs, Paris 2005).

Artykuł Karla Schwarzschilda na temat dopuszczalnej krzywizny przestrzeni nosi tytuł *Über das zulässige Krümmungsmaass des Raumes* i ukazał się w „Vierteljahresschrift der Astronomischen Gesellschaft" 35, 337–347 (1900). Przed Schwarzschildem istnienie skończonego sferycznego wszechświata rozważał już astronom Karl Friedrich Zöllner (1834–1882), jednak nie wspominał on o wariantach złożonych pod względem topologicznym.

W kwestii opisanego w tekście powrotu w formie lustrzanego odbicia zob. również: H.G. Wells, *W czwartym wymiarze* (*Opowieści fantastyczne*, t. 2, Wydawnictwo Literackie, Kraków 1976).

Stała kosmologiczna i ciemna energia

W odniesieniu do stałej kosmologicznej istnieje wiele pierwszorzędnych fachowych artykułów przeglądowych. Prawdziwym klasykiem jest: S. Weinberg, *The cosmological constant problem*, „Reviews of Modern Physics" 61, 1–23 (1989). Na obserwacjach koncentrują się: V. Sahni oraz A. Starobinsky, *The Case for a Positive Cosmological Lambda-term*, „International Journal of Modern Physics" D 9, 373–444 (2000); zob. również: http://arxiv.org/abs/astro-ph/9904398.

W przyszłości planowane są dalsze i dokładniejsze obserwacje supernowych w kosmosie. W tym kontekście warto przede wszystkim wspomnieć o projekcie SNAP. Więcej informacji na jego temat można znaleźć pod tym adresem: http://snap.lbl.gov/.

W kwestii Pauliego i jego kalkulacji wielkości wszechświata nie znam żadnej publikacji, jednak zachowały się cytaty wielu osób, które dyskutowały na ten temat z Paulim; zob. uwagi w publikacji N. Straumanna, http://arxiv.org/abs/astro-ph/0203330, rozdział 2.

W odniesieniu do wnioskowanej z obserwacji przyspieszonej ekspansji wszechświata: obserwacje supernowych i kosmicznego promieniowania tła wskazują mniej lub bardziej dobitnie na istnienie stałej kosmologicznej (lub ciemnej energii), która powoduje przyspieszoną ekspansję wszechświata – pod warunkiem, że interpretacja danych odbywa się na

podstawie modelu Friedmana-Lemaître'a. Nie jest to nieistotne. Z uwagi na to, że równania Einsteina są nieliniowe, możliwe jest imitowanie takiego efektu nawet przez samo wspomnienie o niewielkich niejednorodnościach. Obliczenia są jednak niezwykle skomplikowane, a zatem nie pozostaje nam nic innego jak tylko cierpliwie czekać na odpowiedź, czy ta alternatywna interpretacja poczynionych obserwacji jest rzeczywiście możliwa do zaakceptowania. Pogląd ten został przedstawiony na przykład w: D. Wiltshire, *Dark energy without dark energy*, http://arxiv.org/abs/0712.3984, oraz w: T. Buchert, *Dark energy from structure: a status report*, „General Relativity and Gravitation" 40, 467–527 (2008); zob. również http://arxiv.org/abs/0707.2153 oraz przytoczoną tam literaturę.

Przyszłość wszechświata

Wspomniana w tekście praca, według której w przypadku dominacji stałej kosmologicznej za 100 miliardów lat widzielibyśmy właściwie tylko naszą Drogę Mleczną połączoną z Galaktyką Andromedy, to: K. Nagamine oraz A. Loeb, *Future evolution of nearby large scale structure in a universe dominated by a cosmological constant*, „New Astronomy" 8, 439–448 (2003); zob. również: http://arxiv.org/abs/astro-ph/0204249. Obliczenia prognozujące wchłonięcie Ziemi przez Słońce za około 7,6 miliarda lat można przeanalizować w: K.P. Schröder oraz R.C. Smith, *Distant future of the Sun and Earth revisited*, http://arxiv.org/abs/0801.4031, jak i w: K.-P. Schröder, *Feuriger Weltuntergang*, „Sterne und Weltraum" 47 (lipiec), 42–48 (2008).

Rozważania, według których z pewnymi osobliwości mielibyśmy się zmierzyć już za 10 milionów lat, można prześledzić w: M.P. Dąbrowski i in., *How far is it to a sudden future singularity of pressure?*, „Physical Review" D 75, artykuł nr 123524 (2007); zob. również: http://arxiv.org/abs/0704.1383.

Pochodzący z Wielkiej Brytanii fizyk Freeman Dyson pokusił się wiele lat temu o prognozę tego, jaką przyszłość skrywa dla nas wszechświat.

Zgodnie z jego oceną gwiazdy o niezbyt dużej masie przeistoczą się kiedyś w białe karły, które wystygną po 10^{14} latach (czyli po stu tysiącach miliardów lat). Planety krążące wokół gwiazd pod wpływem innej gwiazdy uwolnią się od swojego centralnego ciała niebieskiego po około 10^{15} latach. Gwiazdy o dużej masie zapadną się, tworząc czarne dziury, które w wyniku działania efektu Hawkinga wyparują po około 10^{65} latach. Zabawy autora z liczbami poprawiają humor, a można je znaleźć w: F.J. Dyson, *Time without end: Physics and biology in an open universe*, „Reviews of Modern Physics" 51, 447–460 (1979). Trudno pozbyć się wrażenia, że przyszło nam żyć w wyjątkowo ciekawym okresie dla wszechświata, a świat, który od dawna znamy, powoli opuszcza stadium stagnacji.

Fluktuacje kwantowe a powstawanie struktur wszechświata

Autorem koncepcji mówiącej o tym, że wszechświat w bardzo wczesnym okresie rozwoju powstał ze stanu próżni, odznaczającego się wysoką symetrią, jest Aleksiej Starobinski. Więcej informacji na ten temat można znaleźć w jego pracy *Spectrum of relict gravitational radiation and the early state of the universe*, „JETP Letters" 30, 682–685 (1979). W publikacji tej przedstawiono ponadto prognozę, zgodnie z którą w omawianym okresie doszło również do powstania fal grawitacyjnych istniejących do dziś. Nie wiemy jednak, czy są one wystarczająco silne, aby udało się je zmierzyć w kosmosie na przykład za pomocą interferometru w ramach planowanej misji LISA. Znaczenie dekoherencji dla fluktuacji pierwotnych omówiono w: C. Kiefer, D. Polarski oraz A.A. Starobinsky, *Quantum-to-classical transition for fluctuations in the early Universe*, „International Journal of Modern Physics" D 7, 455–462 (1998); zob. również: http://arxiv.org/abs/gr-qc/9802003.

Do części 5

Dlaczego grawitacja nie jest klasyczna?

Zwolennikiem wiodącej pozycji redukcjonizmu jest na przykład amerykański laureat Nagrody Nobla Steven Weinberg, przytaczający w swojej książce *Sen o teorii ostatecznej* (Alkazar, Warszawa 1994) wiele argumentów przemawiających za tym stanowiskiem.

Losy Matwieja Bronsteina i opracowanej przez niego fizyki bardzo trafnie opisano w: G.E. Gorelik oraz Y.Y. Frenkel, *Matvei Petrovich Bronstein and Soviet Theoretical Physics in the Thirties* (Birkhäuser, Basel 1994). W kwestii możliwego istnienia najmniejszej długości w zespole skupionym wokół Bronsteina i Landaua powstał również ten wers: „Osły, nie uczeni // Kwantyzacja przestrzeni". (W oryginale: „Die Esel fassen kaum es // Die Quantelung des Raumes" – przyp. tłum.).

Omawiany w tekście eksperyment mający na celu sprawdzenie semiklasycznych równań Einsteina opisano w: D.N. Page oraz C.D. Geilker, *Indirect evidence for quantum gravity*, „Physical Review Letters" 47, 979–982 (1981). W prawdziwym eksperymencie zegar i materiał wybuchowy zostały zastąpione fizykiem, który w zależności od rozpadu źródła poruszał masami.

Jak skwantyzować grawitację?

W odniesieniu do jednostek Plancka: zacytowana praca Plancka na temat nieodwracalnych procesów promieniowania nosi tytuł *Über irreversible Strahlungsvorgänge* i ukazała się w 1899 roku w ramach protokołów z posiedzeń Królewsko-Pruskiej Akademii Nauk w Berlinie: „Sitzungsberichte der königlich-preußischen Akademie der Wissenschaften zu Berlin (physikalisch-mathematische Klasse)", s. 440–480. Już w 1881 roku irlandzki fizyk George Johnstone Stoney (1826–1911) przedstawił jednostki zbliżone do jednostek Plancka. Oczywiście Stoney nie mógł jeszcze wiedzieć o kwancie działania, jednak na podstawie przeprowadzonych eksperymentów mógł on formułować wnioski w kwestii wielkości elementarnego ładunku elektrycznego *e*. Z *e, G* oraz *c* można stworzyć jednostki, które w znacznym stopniu pokrywają się z jednostkami Plancka. Fakt ten jest związany z tym, że tak zwana

stała struktury subtelnej $\alpha = e^2 / \hbar c \approx 1/137$ jest liczbą bezwymiarową. W gruncie rzeczy trzy wielkości wymiarowe (określające „przestrzeń, czas i materię") wydają się wystarczające, aby stworzyć wszystkie jednostki fizyczne (np. wat czy amper).

Oto jednostki Plancka w równaniach:

$$l_{\mathrm{P}} = \sqrt{\frac{\hbar G}{c^3}} \approx 1{,}62 \times 10^{-35}\ \mathrm{m},$$

$$t_{\mathrm{P}} = \frac{l_{\mathrm{P}}}{c} = \sqrt{\frac{\hbar G}{c^5}} \approx 5{,}40 \times 10^{-44}\ \mathrm{s},$$

$$m_{\mathrm{P}} = \frac{\hbar}{l_{\mathrm{P}} c} = \sqrt{\frac{\hbar c}{G}} \approx 2{,}17 \times 10^{-8}\ \mathrm{kg}.$$

Wynikają one z następujących wartości $c \approx 3{,}00 \times 10^8$ m/s, $\hbar \approx 1{,}055 \times 10^{-34}$ Js, $G \approx 6{,}67 \times 10^{-11}$ $\mathrm{Nm^2kg^{-2}}$, przy czym $\hbar = h / 2\pi$, gdzie h to kwant działania. W wymienionych wyżej jednostkach fizyki cząstek elementarnych masa Plancka wynosi około 10^{19} GeV – 10^{19} razy więcej niż proton!

Wspomniany w tekście ból głowy Einsteina znany jest w literaturze jako „problem dziury". W ramach eksperymentu myślowego Einstein wyobraził sobie obszar czasoprzestrzeni pozbawiony materii (czyli „dziurę"), w którym przekształcił metrykę i punkty czasoprzestrzeni niezależnie od siebie i na tej podstawie wywnioskował, że równania pól nie mogły być „ogólnie kowariantne" (czyli nie mogły znajdować się w odrębnych układach współrzędnych o tej samej formie), jeśli szuka się jednoznacznego rozwiązania równań dla wyznaczonego rozkładu materii poza dziurą. Następnie, poprawiając się, napisał w liście do swojego długoletniego przyjaciela Michele Besso, że pod względem fizycznym nic nie jest bardziej *realne* niż ogół czasoprzestrzennych koincydencji

punktów. Ponadto według niego najbardziej naturalnym zachowaniem jest wymaganie od praw, aby nie określały *więcej* niż ogół czasoprzestrzennych koincydencji. Powszechnie znanym przykładem koincydencji czasoprzestrzennej jest system nawigacji satelitarnej GPS, w przypadku którego mamy do czynienia z koincydencjami promieni światła, satelitów i pola grawitacyjnego.

To, że matematyczne mnogości punktowe nie muszą mieć nic wspólnego z rzeczywistością, pokazuje dobitnie paradoks Banacha-Tarskiego, według którego kulę można przedzielić na sześć części, a następnie złożyć z nich dwie kule o wielkości odpowiadającej kuli pierwotnej. (Są to tak zwane niemierzalne mnogości punktowe).

Grawitacje, nieskończoności, całki po trajektoriach

Przedstawioną w tekście procedurę obliczenia całki po trajektoriach w grawitacji kwantowej za pomocą 4-sympleksów określa się również mianem „dynamicznej triangulacji". Jej głównymi pomysłodawcami są Jan Ambjørn, Renate Loll oraz Jerzy Jurkiewicz. Artykuł wprowadzający tych autorów to *The Universe from Scratch*, zamieszczony w „Contemporary Physics" 47, 103–117 (2006); zob. również: http://arxiv.org/abs/hepth/0509010. Procedura ta sprawdza się wyłącznie przy zastosowaniu dodatniej stałej kosmologicznej, co jest zgodne z obserwacjami, jednak jej wartość pozostaje nieokreślona.

Trudno wyobrazić sobie czterowymiarowe obiekty. Wspomniany w tekście 4-sympleks dopiero wtedy nabiera kształtów, gdy obróci się go w rzucie trójwymiarowym (co na papierze oczywiście nie jest możliwe). Odpowiednią wizualizację można znaleźć pod adresem: http://en.wikipedia.org/wiki/Pentachoron.

Kanoniczna grawitacja kwantowa i problem czasu

Wykorzystaną przez Erwina Schrödingera analogię między mechaniką a optyką można znaleźć w jego pracy *Quantisierung als Eigenwertproblem*

(Zweite Mitteilung), „Annalen der Physik" 384, 489–527 (1926). Na stronie 496 pisze on: „Jednak już pierwsza próba sformułowania teorii falowej prowadzi do powstania tak zdumiewających rzeczy, że rodzi się zupełnie inne podejrzenie: *wiemy przecież obecnie, że nasza mechanika klasyczna zawodzi w obliczu bardzo niewielkich wymiarów i bardzo dużych krzywizn toru*. Może niedoskonałość ta jest w pełni analogiczna do niedoskonałości optyki geometrycznej, tzn. «optyki o nieskończenie niewielkiej długości fali», z którą jak wiadomo mamy do czynienia, gdy «przeszkody» lub «szczeliny» nie są dostatecznie duże dla prawdziwej, skończonej długości fali. Może nasza mechanika klasyczna jest w *pełni* analogiczna do optyki geometrycznej, a więc błędna, niezgodna z rzeczywistością (...). Oznacza to zatem, że należy szukać «mechaniki undulacyjnej», a najprostszą drogą do osiągnięcia tego celu jest rozszerzenie analogii Hamiltona na teorię falową.

W odniesieniu do warunków koniecznych: w tekście mówiliśmy o warunku koniecznym Hamiltona jako o centralnym równaniu w kanonicznej koncepcji ogólnej teorii względności. W rzeczywistości istnieją jeszcze trzy inne warunki konieczne, których nie wymieniliśmy, aby nie komplikować sytuacji. Są to „warunki konieczne pędu" lub też „warunki konieczne dyfeomorfizmu", mówiące o tym, że w obrębie jednej przestrzeni trójwymiarowej można wybrać dowolne współrzędne. Ich skwantyzowana forma oznacza, że funkcja falowa nie zmienia się, gdy zostaną wybrane inne współrzędne, co jest logiczne, ponieważ żadna sensowna wielkość fizyczna nie może zależeć od w gruncie rzeczy dowolnych współrzędnych.

W odniesieniu do równania Wheelera-DeWitta: mianem superprzestrzeni określa się właściwie przestrzeń wszystkich trójwymiarowych geometrii, wywodzącą się z przestrzeni wszystkich 3-metryk, gdy dwie metryki, różniące się tylko transformacją współrzędnych, uznaje się za równowartościowe. Ponadto trzeba dodać, że analogia do równania Schrödingera niezależnego od czasu w rozumieniu teorii pola jest trafna, ponieważ funkcja falowa (a z matematycznego

punktu widzenia raczej: funkcjonał falowy) jest opisana w przestrzeni o nieskończenie wielu wymiarach, natomiast tradycyjne równanie Schrödingera w rozumieniu mechaniki kwantowej posiada zaledwie skończoną liczbę stopni swobody (a zatem również skończoną liczbę „cząstek").

W odniesieniu do poszukiwań straconego czasu: procedura odzyskiwania parametru czasu *t* z bezczasowej grawitacji kwantowej jest analogiczna do metody określanej mianem „przybliżenia Borna-Oppenheimera" z zakresu fizyki molekularnej. W tamtym przypadku istnieją dwie różne skale masy (masa ciężkich jąder i lekkich elektronów), które w kontekście grawitacji kwantowej odpowiadają różnicy między masą Plancka a lżejszymi masami. W fizyce molekularnej istnieje już rzecz jasna czas zewnętrzny, jednak zastosowany formalizm jest w dużej mierze analogiczny.

Grawitacja kwantowa z pętlami

Cytat Penrose'a można znaleźć w jego pracy *Angular momentum: an approach to combinatorial space-time*, zamieszczonej w *Quantum Theory and Beyond* pod redakcją T. Bastina (Cambridge University Press, Cambridge 1971), s. 151–180. Nazwa wspomnianego w tekście „parametru Barbero-Immirziego" pochodzi od nazwisk dwóch fizyków: Giorgia Immirziego (Perugia) oraz Fernanda Barbero (Madryt).

W teorii pętlowej dyskretna natura przestrzeni daje o sobie znać w najmniejszych skalach. Na myśl automatycznie przychodzi odwieczna dyskusja między zwolennikami dyskretnej i ciągłej przestrzeni. W tym kontekście warto przytoczyć wypowiedź Wolfganga Pauliego. Tak napisał on w liście z 16 października 1957 roku do Markusa Fierza: „Nie mam wątpliwości co do tego, że klasyczna fizyka pola «wywodzi się» w linii prostej ze stoicyzmu i była rozwijana w sposób *ciągły* oraz zgodny z renesansowymi i siedemnastowiecznymi ideami (...). Skoro problem syntezy teorii kwantowej i ogólnej teorii względności (i w ogóle kwantyzacji pola) pozostaje otwarty, to stary (antyczny) konflikt między atomistami a stoikami nadal trwa".

List ten można znaleźć w książce: K. von Meyenn, *Wissenschaftlicher Briefwechsel*, t. IV, cz. IV (Springer, Berlin 2005), s. 571.

Teoria strun

W odniesieniu do modelu strunowego oddziaływań silnych: pierwotny model strunowy oddziaływań silnych został obalony doświadczalnie. Mimo tego faktu obecnie ponownie mówi się o takich modelach, w których dawne sprzeczności są unieważniane w wyniku przyjęcia hipotezy o piątym wymiarze; zob. na przykład: J. Maldacena, *Into the fifth dimension*, „Nature" 423, 695–696 (2003). Celem tej hipotezy jest uproszczenie obliczeń chromodynamiki kwantowej.

W odniesieniu do konieczności istnienia grawitacji w teorii strun: w rzeczywistości grawiton występuje tylko w spektrum zamkniętych strun. Z uwagi jednak na to, że teoria otwartych strun obejmuje również zamknięte struny w formie wirtualnej, prognozuje ona istnienie grawitonu, a tym samym także grawitacji.

W odniesieniu do pięciu różnych teorii strun: pięć wymienionych w tekście teorii strun różni się pod względem sposobu implementacji supersymetrii. Wszystkie te teorie mają swoje nazwy, które przynajmniej powinniśmy wymienić: teoria typu I, teoria typu IIA, teoria typu IIB, teoria heterotyczna typu O(32) oraz teoria heterotyczna typu $E_8 \times E_8$. Jedyną teorią odnoszącą się do strun otwartych jest teoria typu I.

W odniesieniu do m-teorii i jedenastego wymiaru: 11 to maksymalny wymiar w przypadku tak zwanej supergrawitacji. Pod tym pojęciem należy rozumieć kwantową teorię pola, wprowadzającą grawitację do teorii supersymetrii. W latach siedemdziesiątych XX wieku supergrawitacja była traktowana jako teoria zunifikowana, uwzględniająca grawitację. Przez długi czas supergrawitacja w jedenastu wymiarach była postrzegana przez fizyków teoretycznych, zajmujących się teorią strun, jako dziwactwo. Powodem tego była teoria strun, która prognozowała istnienie tylko dziesięciu wymiarów. Pojawienie się m-teorii i jej jedenastu wymiarów zmieniło wszystko. Jedenastowymiarowa supergrawitacja występuje

w przypadku granicznym przy niewielkich energiach. (Dla fizyka pracującego w Kolonii liczba jedenaście ma oczywiście szczególne znaczenie).

Nie jest żadnym zaskoczeniem, że teorie takie jak teoria strun, których nie przetestowano jeszcze doświadczalnie, spotykają się ze skrajnymi ocenami. Zalety tej teorii przedstawia na przykład B. Greene w książce *Piękno wszechświata* (zob. spis literatury). Pozycję skrajnie przeciwną zajął P. Woit w książce *Not Even Wrong: The Failure of String Theory and the Continuing Challenge to Unify the Laws of Physics* (Jonathan Cape, London 2006).

Wszechświat na branie

Omawiane w tekście modele o większej liczbie wymiarów bardzo szczegółowo i zrozumiale opisała w swojej książce Lisa Randall (zob. spis literatury). Czasoprzestrzeń między obiema branami w modelu Randall i Sundruma (które określa się angielskim słowem „bulk", co oznacza mniej więcej tyle co „część zasadnicza") jest częścią czasoprzestrzeni anty-de Sittera, czyli rozwiązania równań Einsteina o ujemnej stałej kosmologicznej (zob. część 4).

Randall i Sundrum opracowali również drugi model, w którym znajduje się tylko jedna brana i piąty wymiar jest nieskończenie wielki. Jest to możliwe i nie jest sprzeczne z obserwacjami, pod warunkiem że grawitacja jest odpowiednio zakrzywiona (z angielskiego: *warped*).

Czy w czarnych dziurach gubią się informacje?

Względy, którymi kierował się Hawking, podejmując decyzję o zmianie zdania, opisał on w artykule *Information loss in black holes*, opublikowanym w „Physical Review" D 72, artykuł nr 084013 (2005); podobna wersja jest dostępna nieodpłatnie pod adresem: http://arxiv.org/abs/hep-th/0507171. Na właściwość czarnej dziury, pozwalającą tworzyć wraz ze swoim otoczeniem oddziałujący otwarty

układ kwantowy, szczególną uwagę zwraca H.D. Zeh w swojej pracy pod tytułem *Where has all the information gone?*, ogłoszonej w „Physics Letters" A 347, 1–7 (2005); podobna wersja jest dostępna nieodpłatnie pod adresem: http://arxiv.org/abs/gr-qc/0507051.

Pierwotne czarne dziury

Podczas wspomnianych w tekście poszukiwań pierwotnych czarnych dziur znajdujemy tylko górne granice. Obliczono, że na każdy parsek sześcienny mogłoby przypadać maksymalnie około 10 000 pierwotnych dziur o masie oscylującej wokół 10^{15} grama. Odpowiada to wkładowi do kosmologicznego parametru omega (część 4) na poziomie nieprzekraczającym wartości 10^{-8}. Dziury takie nie są zatem istotną częścią ciemnej materii. Nie jest jednak wykluczone, że cięższe minidziury, na przykład te, których masa oscyluje wokół 10^{20} grama, dostarczą nam istotnych informacji na temat ciemnej materii w galaktykach. Dopóki nie wiemy, z jak wieloma takimi obiektami mamy do czynienia, sprawa ta pozostaje nadal otwarta. W kwestii obserwacji bezpośrednich zjawisk odparowywania wyznaczono natomiast górną granicę na poziomie około miliona zdarzeń na parsek sześcienny i rok. Podsumowanie aktualnego stanu wiedzy można znaleźć w artykule *Primordial Black Holes: Do They Exist and Are They Useful?* autorstwa B.J. Carra, dostępnym pod adresem: http://arxiv.org/abs/astro-ph/0511743.

Do części 6

Grawitacja kwantowa i kosmologia

W odniesieniu do wielkiego zatrzymania: wspomnianą w tekście dyskusję na temat wielkiego zatrzymania można znaleźć w: A.Y. Kamenshchik, C. Kiefer oraz B. Sandhöfer, *Quantum cosmology with a big-brake singularity*, „Physical Review" D 76, artykuł nr 064032 (2007); podobna

wersja jest dostępna nieodpłatnie pod adresem: http://arxiv.org/abs/0705.1688.

W odniesieniu do propozycji Hartle'a i Hawkinga: Hawking po raz pierwszy przedstawił „warunek braku granic" na konferencji w Watykanie w 1982 roku. W sprawozdaniach z konferencji, opublikowanych w 48. tomie „Pontificia Academiae Scientarium Scripta Varia", znajduje się wstęp napisany przez papieża, według którego: „Pismo Święte chce ponadto przekazać ludziom, że świat (...) został stworzony na służbę człowiekowi i na cześć Boga. Wszelka inna teoria o powstaniu świata jest sprzeczna z intencjami Biblii (...).

Żadna naukowa hipoteza na temat pochodzenia świata, jak na przykład ta o prymitywnym atomie (...), nie potrafi rozwiązać problemu powstania wszechświata. Cała nauka nie jest w stanie poradzić sobie z tą kwestią". Hawking powiedział później, że jest zadowolony, ponieważ papież najprawdopodobniej nie zrozumiał jego wykładu, którego konkluzja była podobna do tego, co w swojej krytyce chciał przekazać papież.

Wspomniany w tekście i przedstawiony na rysunku 6.2 instanton Hartle'a-Hawkinga był już wcześniej tematem rozważań Aleksandra Wilenkina (1982), jednak w innym kontekście.

W odniesieniu do warunku tunelowego: ciekawe spostrzeżenie na ten temat można przeczytać w: H. Blumenberg, *Die Vollzähligkeit der Sterne* (Suhrkamp, Frankfurt am Main 1997), s. 46: „W odpowiedzi na pytanie Leibniza o to, dlaczego istnieje raczej coś niż nic, jeden z kosmologów stwierdził abstrakcyjnie: «Ponieważ nic jest niestabilne»". Blumenberg skomentował te słowa, mówiąc ironicznie, że twierdzenie to może nie być błędne, gdyż o nicości można przecież powiedzieć wszystko. Nie oznacza to jednak, że twierdzenie owo jest prawdziwe (w rozumieniu fizycznym), ale może być po prostu pozbawione sensu.

Jeden czy wiele światów?

W tekście wspomniano, że właściwie wszystkie omawiane obecnie ujęcia grawitacji kwantowej są liniowe, a zatem prowadzą do obrazu świata

kwantowego według Everetta. Jednym z wyjątków jest propozycja Rogera Penrose'a, rozważającego możliwość wystąpienia kolapsu spowodowanego grawitacją (zob. na przykład: R. Penrose, *Nowy umysł cesarza: o komputerach, umyśle i prawach fizyki*, Wydawnictwo Naukowe PWN, Warszawa 1995). Jest to bardzo interesująca ewentualność, która niestety współcześnie nie została rozwinięta ani poparta eksperymentem.

Powstanie kierunku upływu czasu

W odniesieniu do nieprawdopodobieństwa naszego wszechświata: w tekście napisano, że istnienie naszego wszechświata jest w ogromnym stopniu nieprawdopodobne, ponieważ stan początkowy musiał być niezwykle jednorodny. Z jakim stopniem nieprawdopodobieństwa mamy jednak do czynienia? Zgodnie z oceną Rogera Penrose'a entropia obserwowanego wszechświata przyjmuje swoją wartość maksymalną, gdy cała masa gromadzi się w czarnej dziurze. Ze wzoru na entropię czarnej dziury (porównaj z częścią 3) wynika zatem entropia o wartości $S_{max} \approx 10^{123} k_B$. Entropia naszego wszechświata jest zdominowana przez kosmiczne promieniowanie tła i wynosi $S \approx 10^{88} k_B$. Prawdopodobieństwo dla naszego wszechświata można obliczyć ze wzoru Einsteina z 1905 roku, który odwraca odwrócony związek Boltzmanna „Entropia jest logarytmem prawdopodobieństwa". Zgodnie z powyższym:

$$\frac{\exp\left(\frac{S}{k_\mathrm{B}}\right)}{\exp\left(\frac{S_\mathrm{max}}{k_\mathrm{B}}\right)} \sim \frac{\exp(10^{88})}{\exp(10^{123})} \approx \exp\left(-10^{123}\right)$$

Prawdopodobieństwo jest więc niewyobrażalnie niskie – jest to liczba z gatunku tych 0,000 ... 1, przy czym 1 znajduje się na 10^{123} miejscu po przecinku. Bardziej prawdopodobne byłoby już nawet założenie, że cały

wszechświat powstał w jednym momencie wraz z wszystkimi ludźmi i ich wspomnieniami w ramach niezwykle nieprawdopodobnej fluktuacji (wahania w rozumieniu Boltzmanna, porównaj z częścią 3). Niska wartość prawdopodobieństwa nie bierze się zatem znikąd.

W odniesieniu do odwrócenia strzałki czasu: możliwość odwrócenia strzałki czasu podczas przejścia od fazy ekspansji do fazy kurczenia się wszechświata wskazał już astrofizyk Thomas Gold w 1962 roku. Przy tym przedstawił on jednak wyłącznie klasyczne i prawdopodobnie niewystarczające argumenty. W 1985 roku Stephen Hawking, zainspirowany warunkiem braku granic, doszedł do tego samego wniosku, choć później, po zgłoszeniu zastrzeżeń przez Dona Page'a, odwołał swoje słowa ze względu na argumenty semiklasyczne. Ciągłe analizy tego problemu w ramach kosmologii kwantowej prowadzą jednak do pierwotnego wniosku; zob. argumentacja w tekście oraz szczegółowe badania w: C. Kiefer oraz H.D. Zeh, *Arrow of time in a recollapsing quantum universe*, „Physical Review" D 51, 4145–4153 (1995). W pozycji tej przedstawiono również konsekwencje, jakie wynikają z tego dla czarnej dziury, która znajduje się w obrębie wszechświata kwantowego: nie istnieje horyzont zdarzeń ani żadna osobliwość w jej wnętrzu. W fazie kurczenia się wszechświata czarne dziury stają się formalnie białymi dziurami, które postrzegane są jednak przez obserwatorów we wszechświecie jako czarne dziury.

Zasada antropiczna

Brandon Carter wprowadził zasadę antropiczną w bardzo ciekawym artykule pod tytułem *Large number coincidences and the anthropic principle in cosmology*, zamieszczonym w książce *Confrontation of Cosmological Theories with Observational Data* pod redakcją M.S. Longaira (D. Reidel, Dordrecht 1974), s. 291–298. Aktualne informacje na temat zasady antropicznej oraz pojęcia multiświata można znaleźć w książce *Universe or Multiverse?* pod redakcją B. Carra (Cambridge University Press, Cambridge 2007).

O zasadzie antropicznej i zasadzie przeciętności pisze również A. Wilenkin w *Many Words in One* (Farrar, Straus and Giroux, New York 2007). Książka zawiera także szczegółowy opis wspomnianego wyżej warunku tunelowego w kosmologii kwantowej.

Do części 7

W tekście nawiązaliśmy do *Teodycei* Leibniza, która uzasadnia, dlaczego nasz świat jest najlepszym ze wszystkich możliwych światów. W części 1 możemy przeczytać: „Zatem rozmiary znanej nam części wszechświata giną niemal w niebycie, gdy porównamy je z tym, czego nie znamy, a co jednak mamy powód zakładać. Podobnie wszelkie zło, które można nam przedstawić jako zarzut, należy jedynie do tego prawie niebytu. Możliwe, że wszelkie zło jest także tylko prawie niebytem w porównaniu z istniejącym we wszechświecie dobrem" (G.W. Leibniz, *Teodycea*, wg przekładu Małgorzaty Frankiewicz, 2001). O ile mniejszy jest ten prawie niebyt w kosmosie kwantowym, w którym się znajdujemy!

Powstało oczywiście wiele publikacji traktujących o tematach omawianych w tej książce. Cytaty Plancka można znaleźć w: M. Planck, *Vorträge und Erinnerungen* (Wissenschaftliche Buchgesellschaft, Darmstadt 1965), s. 228–245. Godny polecenia artykuł Einsteina pod tytułem *Autobiographisches* przedrukowano w książce *Albert Einstein als Philosoph und Naturforscher* pod redakcją P.A. Schilppa (W. Kohlkammer-Verlag, Stuttgart 1955). Zacytowane słowa Wittgensteina można znaleźć na stronie 421 książki *Wiener Ausgabe: „The Big Typescript"* (Springer, Wien 2000). W kwestii dylematu wolnej woli warto sięgnąć po *Das Handwerk der Freiheit* (S. Fischer, Frankfurt am Main 2003) autorstwa P. Bieriego.

Literatura uzupełniająca

Literatura ogólna

Barbour J.B., *The End of Time*, Weidenfels & Nicolson, London 1999.

Bruß D., *Quanteninformation*, S. Fischer, Frankfurt am Main 2003.

Davies P.C.W. oraz Brown J. (red.), *Superstrings – A Theory of Everything?*, Cambridge University Press, Cambridge 1988.

Einstein A., *Autobiographisches*, w: P.A. Schilpp (red.), *Albert Einstein als Philosoph und Naturforscher*, Vieweg, Braunschweig 1983.

Eisenhardt P., *Der Webstuhl der Zeit*, Rowohlt Taschenbuch Verlag, Reinbek 2006.

Fölsing A., *Albert Einstein. Eine Biographie*, Suhrkamp, Frankfurt am Main 1995.

Genz H., *Wie die Zeit in die Welt kam*, Carl Hanser Verlag, München 1996.

Giulini D., *Spezielle Relativitätstheorie*, S. Fischer, Frankfurt am Main 2004.

Greene B., *Piękno wszechświata*, Prószyński i S-ka, Warszawa 2001.

Greene B., *The Fabric of the Cosmos: Space, Time, and the Texture of Reality*, Vintage Books, New York 2004.

Gribbin J., *Kotki Schrödingera, czyli poszukiwanie rzeczywistości*, Zysk i S-ka, Poznań 1999.

Hasinger G., *Das Schicksal des Universums: Eine Reise vom Anfang zum Ende*, C.H. Beck, München 2007.

Hawking S., *Krótka historia czasu: od wielkiego wybuchu do czarnych dziur*, Alfa, Warszawa 1990.

Hawking S. oraz Penrose R., *Natura czasu i przestrzeni*, Zysk i S-ka, Poznań 1996.

Kiefer C., *Gravitation*, S. Fischer, Frankfurt am Main 2003.

Kiefer C., *Quantentheorie*, wydanie trzecie, S. Fischer, Frankfurt am Main 2004.

Kippenhahn R., *Licht vom Rande der Welt*, Piper, München 1987.

Randall L., *Ukryte wymiary wszechświata*, Prószyński i S-ka, Warszawa 2011.

Schwinger J., *Einstein's Legacy*, Scientific American Library, New York 1986.

Smolin L., *Trzy drogi do kwantowej grawitacji*, CiS, Warszawa 2001.

Tegmark M. oraz Wheeler J.A., *Stulecie kwantowych zagadek*, „Świat Nauki”, kwiecień 2001, s. 54–62.

Thorne K., *Czarne dziury i krzywizny czasu: zdumiewające dziedzictwo Einsteina*, Prószyński i S-ka, Warszawa 2004.

Vaas R., *Tunnel durch Raum und Zeit*, Kosmos, Stuttgart 2005.

Wheeler J.A., *A Journey into Gravity and Spacetime*, Scientific American Library, New York 1990.

Zeh H.D., *Entropie*, S. Fischer, Frankfurt am Main 2005.

Literatura fachowa

d’Espagnat B., *Veiled Reality*, Westview Press, Boulder 2003.

Giulini D., Kiefer C. oraz Lämmerzahl C. (red.), *Quantum Gravity: From Theory to Experimental Search*, Springer, Berlin 2003.

Harrison E., *Cosmology*, wydanie drugie, Cambridge University Press, Cambridge 2000.

Hartle J.B., *Grawitacja. Wprowadzenie do ogólnej teorii względności Einsteina*, Wydawnictwa Uniwersytetu Warszawskiego, Warszawa 2010.

Joos E., Zeh H.D., Kiefer C., Giulini D., Kupsch J. oraz Stamatescu I. O., *Decoherence and the Appearance of a Classical World in Quantum Theory*, wydanie drugie, Springer, Berlin 2003.

Kiefer C., *Quantum Gravity*, wydanie drugie, Oxford University Press, Oxford 2007.

Rovelli C., *Quantum Gravity*, Cambridge University Press, Cambridge 2004.

Zeh H.D., *The Physical Basis of the Direction of Time*, wydanie piąte, Springer, Berlin 2007.

Zwiebach B., *A First Course in String Theory*, Cambridge University Press, Cambridge 2004.